The chain rule for functionals with applications to functions of moments

by

Christopher S. Withers
Applied Mathematics Group
Industrial Research Limited
Lower Hutt, NEW ZEALAND

Saralees Nadarajah
School of Mathematics
University of Manchester
Manchester M13 9PL, UK

Abstract: The chain rule for derivatives of a function of a function is extended to a function of a statistical functional, and applied to obtain approximations to the cumulants, distribution and quantiles of functions of sample moments, and so to obtain third order confidence intervals and estimates of reduced bias for functions of moments. As an example we give the distribution of the standardized skewness for a normal sample to magnitude $O(n^{-2})$, where n is the sample size.

AMS 2000 subject classification: Primary 62E20; Secondary 62G30.

Keywords and phrases: Asymptotic expansions; Bias; Chain rule; Confidence intervals; Cumulants; Moments; Nonparametric.

1 Introduction

The derivatives introduced by von Mises (1947) and their subsequent versions have wide ranging applications in statistics. Two prominent application areas are the construction of nonparametric confidence intervals and analytic bias reduction.

Suppose we want to construct a nonparametric confidence interval of level $\alpha + O(n^{-3/2})$ for a smooth functional $T(F)$ based on $\widehat{F}$ say, the sample or empirical distribution for a sample of size n from F. Withers (1983) showed that the limit can be given in terms of integrals of products of von Mises derivatives evaluated at $\widehat{F}$. First one Studentizes using the asymptotic variance of $n^{1/2}\{T(\widehat{F}) - T(F)\}$, $a_{21}(F) = [1^2]_T = \int_{-\infty}^{\infty} T_F(x_1)^2 dF(x_1)$, where $T_F(x)$ is the first derivative or influence function of $T(F)$.

Similarly, it is known that for smooth $T(F)$, an estimate of $T(F)$ of bias $O(n^{-2})$ is $T(\widehat{F}) - n^{-1}a_{11}(\widehat{F})$, where $a_{11}(F) = [11]_T = \int_{-\infty}^{\infty} T_F(x_1, x_1) dF(x_1)$ and $T_F(x_1, x_2)$ is the second derivative of $T(F)$. For convenience we refer to these integrals of products of derivatives like $[1^2]_T$ and $[11]_T$ as *bracket functions.*

Other recent application areas of von Mises derivatives include: least squares support vector regression filtering methods, bootstrapping, functional principal components analysis, linearization and composite estimation, dimension reduction, quantile regressions,

machine learning, cusum statistics, methods of sieves and penalization, change point estimation, Hadamard differentiability, change-of-variance function, measuring and testing dependence by correlation of distances, empirical finite-time ruin probabilities (Loisel *et al.*, 2009), nonparametric maximum likelihood estimators (Nickl, 2007), estimating mean dimensionality of analysis of variance decompositions, monotonicity of information in the Central Limit Theorem, generalizations of the Anderson-Darling statistic, M-estimation, U-statistics (Volodko, 2011), information criteria in model selection, goodness-of-fit tests for kernel regression, empirical Bayes estimation, and estimation of Kendall's tau.

The aim of this paper is to develop tools for extending the use of von Mises derivatives. In Section 2, we extend Faa di Bruno's chain rule for the derivative of a function of a univariate function to functions of a multivariate function and show how it can be applied to a function of a function of F, say $T(F) = g(S(F))$, where $g : \mathbb{R}^a \to \mathbb{R}$ is a smooth function and $S(F)$ a smooth functional.

In Section 3, we apply it to obtain derivatives and bracket functions for powers, products, quotients, standardized and Studentized functionals.

Section 4 gives the general derivative for a moment and applies previous results to obtain expansions up to $O(n^{-2})$ for the distribution and quantiles of functions of sample moments. As an example we give the distribution of the standardized skewness for a normal sample to magnitude $O(n^{-2})$, where n is the sample size. Also we give confidence intervals and bias reduction methods for functions of moments.

Some of the results in the paper follow easily from Withers (1983, 1987), see Theorems 3.1 to 3.3. But these results are not the main contributions of this paper. The main contributions are: 1) the tools developed to compute von Mises type derivatives, see Theorems 2.1 and 2.2; 2) their applications to obtain bracket functions for general functionals, see Examples 3.1 to 3.4 and Appendix A. The functionals considered by these examples include $T(F) = g(S(F))$, where g is a univariate function, $T(F) = S_1(F)S_2(F)$, a product of two functionals, Studentized forms of $T(F)$ and $T(F) = U(F)g(S(F))$, where $S(F)$ is real valued; 3) also the applications of Theorems 2.1 and 2.2 to obtain derivatives of central moments, see Theorem 4.1, Corollary 4.1, Corollary 4.2 and Appendix B.

Fisher and Wishart gave unbiased estimates only for cumulants and their products: see, for example, Stuart and Ord (1987). Our two methods for bias reduction apply to any smooth functional - and our second estimate reduces to their results for the cases they consider. Also our method does not need to use unbiased estimates of cumulants to reduce the bias of functions of cumulants.

Analogous to Fisher's tables for his k-statistics and their cumulants, Appendix B gives the terms needed for bias reduction of any smooth function of one or more moments.

2 Chain rules for functions and functionals

Let s and g be real functions on $\mathbb{R}$ with finite derivatives. Comtet (1974, page 137) gives Faa di Bruno's chain rule for the rth derivative of

$$t(x) = g\left(s(x)\right)$$

for $r = 1, 2, \ldots$ in the form

$$t^{(r)}(x) = \sum_{h=1}^{r} g^{(h)}\left(s(x)\right) B_{rh}(\mathbf{s}) \tag{2.1}$$

evaluated at $\mathbf{s} = (s_1, s_2, \ldots)$, $s_i = s^{(i)}(x)$, where B_{rh} is the *partial exponential Bell polynomial* defined by the coefficients in the formal expansion in powers of real ε,

$$\left(\sum_{i=1}^{\infty} \varepsilon^i s_i/i!\right)^j /j! = \sum_{r=j}^{\infty} \varepsilon^r B_{rj}(\mathbf{s})/r!$$

for $j \geq 0$. Comtet (1974) shows they are given by

$$B_{rj}(\mathbf{s})/j! = \sum_{n \text{ in } \mathbb{N}^r} \left\{ \frac{s_1^{n_1} \cdots s_r^{n_r}}{n_1! \cdots n_r!} : n_1 + \cdots + n_r = j,\ 1 \cdot n_1 + \cdots + r \cdot n_r = r \right\},$$

where $\mathbb{N} = \{0, 1, 2, \ldots\}$. Comtet (1974, page 307) tables them for $r \leq 12$. For example,

$$B_{r1}(\mathbf{s}) = s_r,\ B_{rr}(\mathbf{s}) = s_1^r, \tag{2.2}$$
$$B_{32}(\mathbf{s}) = 3s_1s_2,\ B_{42}(\mathbf{s}) = 4s_1s_3 + 3s_2^2,\ B_{43}(\mathbf{s}) = 6s_1^2s_2. \tag{2.3}$$

Theorem 2.1 provides an extension of (2.1) to the case $s : \mathbb{R}^a \to \mathbb{R}^b$ and $g : \mathbb{R}^b \to \mathbb{R}$.

Theorem 2.1 *Define the partial derivatives*

$$\begin{aligned}
&t_{\cdot j_1 \cdots j_r}(x) = \partial^r t(x)/\partial x_{j_1} \cdots \partial x_{j_r},\\
&s_{i \cdot j_1 \cdots j_r} = s_{i \cdot j_1 \cdots j_r}(x) = \partial^r s_i(x)/\partial x_{j_1} \cdots \partial x_{j_r},\\
&g_{\cdot j_1 \cdots j_r}(s) = \partial^r g(s)/\partial s_{j_1} \cdots \partial s_{j_r}.
\end{aligned}$$

The extension of (2.1) is

$$t_{\cdot 1 \cdots r}(x) = \sum_{h=1}^{r} g_{\cdot i_1 \cdots i_h}\left(s(x)\right) B_r^{i_1 \cdots i_h}(\mathbf{s}). \tag{2.4}$$

In (2.4) and throughout, we use the tensor sum convention that repeated indices $i_1, i_2, \ldots$ are implicitly summed over their range ($1, \ldots, b$ in the case of (2.4)).

Note that $B_r^{i_1 \cdots i_h}(\mathbf{s})$ can be written down on sight from B_{rh}. Some particular cases of $B_r^{i_1 \cdots i_h}(\mathbf{s})$ can be obtained from (2.2) and (2.3):

$$\begin{aligned}
&B_r^{i}(\mathbf{s}) = s_{i \cdot 1 \cdots r},\ B_r^{i_1 \cdots i_r}(\mathbf{s}) = s_{i_1 \cdot 1} \cdots s_{i_r \cdot r},\\
&B_3^{i_1 i_2}(\mathbf{s}) = \sum^{3} s_{i_1 \cdot 1} s_{i_2 \cdot 23},\\
&B_4^{i_1 i_2}(\mathbf{s}) = \sum^{4} s_{i_1 \cdot 1} s_{i_2 \cdot 234} + \sum^{3} s_{i_1 \cdot 12} s_{i_2 \cdot 34},\ B_4^{i_1 i_2 i_3}(\mathbf{s}) = \sum^{6} s_{i_1 \cdot 1} s_{i_2 \cdot 2} s_{i_3 \cdot 34},
\end{aligned}$$

where

$$\sum^{r} h_{1 \cdots r} = h_{1 \cdots r} + h_{2 \cdots r 1} + \cdots + h_{r 1 \cdots r-1}.$$

A form of the multivariate chain rule (2.4) was given in Withers (1984).

Let $\mathcal{F}$ be a convex set of probability measures on a measurable space (Ω, A). Suppose for $x \in \Omega$, that δ_x lies in $\mathcal{F}$, where δ_x is the measure putting mass 1 at x and 0 elsewhere. Let x, $\{x_i\}$ be points in Ω. Let F lie in $\mathcal{F}$, and $T : F \to \mathbb{R}$ be some functional. Define the rth derivative of $T(F)$ at $(x_1, \ldots, x_r)$,

$$T_{\cdot 1 \cdots r} = T_F\left(x_1 \cdots x_r\right) = T_F^{(r)}\left(x_1, \ldots, x_r\right),$$

as in Withers (1983). The only derivative we need give here is the first, also known as the *influence function*:

$$T_{\cdot 1}\left(x_1\right) = T_F\left(x_1\right) = \lim_{\epsilon \downarrow 0}\left\{T\left((1-\epsilon)F + \epsilon\delta_{x_1}\right) - T(F)\right\}/\epsilon.$$

For example, $T(F) = \int_{-\infty}^{\infty} g(x)dF(x)$ has first derivative $T_{\cdot 1} = T_{\cdot 1}(x) = g(x) - T(F)$. The results stated in Withers (1983) for $\Omega = \mathbb{R}^s$ generalize immediately to general Ω. In particular, the rule (2.11) for the derivative of the rth derivative may be stated as

$$\left(T_{\cdot 1 \cdots r}\right)_{r+1} = T_{\cdot 1 \cdots r+1} - \sum_{i=1}^{r}\left[T_{\cdot 1 \cdots r+1}\right]_i, \tag{2.5}$$

where $[T_{\cdot 1 \cdots r+1}]_i = T_{\cdot 1 \cdots r+1}$ with the ith argument dropped. So,

$$\begin{aligned}
\left(T_{\cdot 1}\right)_2 &= T_{\cdot 12} - T_{\cdot 2}, \\
\left(T_{\cdot 12}\right)_3 &= T_{\cdot 123} - T_{\cdot 23} - T_{\cdot 13}.
\end{aligned}$$

In this way higher derivatives may be calculated from successive first derivatives. For example, the second derivative of $\int_{-\infty}^{\infty} g(x)dF(x)$ is zero. Now suppose for some function $g : \mathbb{R}^b \to \mathbb{R}$,

$$T(F) = g\left(S(F)\right) \text{ where } S(F) \text{ is a real functional in } \mathbb{R}^d. \tag{2.6}$$

Applying (2.5) gives

$$T_{\cdot 1} = g_{\cdot i} S_{i \cdot 1}, \tag{2.7}$$

$$T_{\cdot 12} = g_{\cdot i} S_{i \cdot 12} + g_{\cdot ij} S_{i \cdot 1} S_{j \cdot 2}, \tag{2.8}$$

$$T_{\cdot 123} = g_{\cdot i} S_{i \cdot 123} + g_{\cdot ij} \sum^{3} S_{i \cdot 1} S_{j \cdot 23} + g_{\cdot ijk} S_{i \cdot 1} S_{j \cdot 2} S_{k \cdot 3}, \tag{2.9}$$

and so on, where $S_{a \cdot 12 \cdots}$ is the rth derivative of $S_a(F)$. Despite the fact that by (2.5) the derivative of a derivative is *not* a second derivative, the expressions (2.7)-(2.9) are precisely those for the derivatives of a function of a vector function of a vector given in (2.4). That is,

$$T_{\cdot 1 \cdots r} = \sum_{h=1}^{r} g_{\cdot i_1 \cdots i_h}\left(S(F)\right) B_r^{i_1 \cdots i_h}(\mathbf{S}), \tag{2.10}$$

where $\mathbf{S} = (S_1, S_2, \ldots)$, $S_i = S^{(i)}(F)$. A proof that (2.10) holds for general r follows using (2.5) and induction. The result can be formally stated as follows.

Theorem 2.2 *If (2.6) holds, $T_{\cdot 1\cdots r}$ is given by the chain rule for*

$$T_{\cdot 1\cdots r} = T_{\cdot 1\cdots r}(x) = \partial^r T(x)/\partial x_1 \cdots \partial x_r$$

when $T(x) = g(S(x))$ for $S(x) : \mathbb{R}^r \to \mathbb{R}^d$ with $S_{i\cdot 1\cdots r} = S_{i\cdot 1\cdots r}(x) = \partial^r S_i(x)/\partial x_1 \cdots \partial x_r$ re-interpreted as $S_{iF}(x_1 \cdots x_r)$ and $S(x)$ as $S(F)$. So,

$$T_{\cdot 1\cdots r} = \sum_{h=1}^{r} g_{\cdot i_1\cdots i_h} \sum_{n} \sum_{\Pi}^{m(n)} S_{i_1\cdot \Pi_1} \cdots S_{i_h\cdot \Pi_h}, \tag{2.11}$$

where $g_{\cdot i_1\cdots i_h} = g_{\cdot i_1\cdots i_h}(y) = \partial_{i_1} \cdots \partial_{i_h} g(y)$ at $y = S(F)$ for $\partial_i = \partial/\partial y_i$, $\sum_n$ sums over $n = (n_1 \cdots n_r) \in \mathbb{N}^r$ satisfying $\sum_{i=1}^r n_i = k$, $\sum_{i=1}^r i n_i = r$, $m(n) = r!/\prod_{i=1}^r i!^{n_i} n_i!$, the partition function, and $\sum_{\Pi}^{m(n)}$ sums over all partitions $(\Pi_1 \cdots \Pi_k)$ of $(1 \cdots r)$ with i Π's of length n_i.

Corollary 2.1 applies Theorem 2.2 to obtain the next two derivatives.

Corollary 2.1 *We*

$$\begin{aligned}
T_{\cdot 1234} &= g_{\cdot i} S_{i\cdot 1234} + g_{\cdot i_1 i_2} \left(\sum^{4} S_{i_1\cdot 1} S_{i_2\cdot 234} + \sum^{3} S_{i_1\cdot 12} S_{i_2\cdot 34} \right) \\
&\quad + g_{\cdot i_1 i_2 i_3} \sum^{6} S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 34} + g_{\cdot i_1\cdots i_4} S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 3} S_{i_4\cdot 4}, \\
T_{\cdot 12345} &= g_{\cdot i} S_{i\cdot 1\cdots 5} + g_{\cdot i_1 i_2} \left(\sum^{4} S_{i_1\cdot 1} S_{i_2\cdot 2345} + \sum^{10} S_{i_1\cdot 12} S_{i_2\cdot 345} \right) \\
&\quad + g_{\cdot i_1 i_2 i_3} \left(\sum^{10} S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 345} + \sum^{15} S_{i_1\cdot 1} S_{i_2\cdot 23} S_{i_3\cdot 45} \right) \\
&\quad + g_{\cdot i_1\cdots i_4} \sum^{10} S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 3} S_{i_4\cdot 45} + g_{\cdot i_1\cdots i_5} S_{i_1\cdot 1} \cdots S_{i_5\cdot 5},
\end{aligned}$$

so

$$\begin{aligned}
\left[1^k\right]_T &= g_{\cdot i_1} \cdots g_{\cdot i_k} \left[S_{i_1\cdot 1} \cdots S_{i_k\cdot 1}\right], \\
[11]_T &= g_{\cdot i}[11]_{S_i} + g_{\cdot i} g_{\cdot j} \left[S_{i\cdot 1} S_{j\cdot 1}\right], \\
[1,2,12]_T &= g_{\cdot i_1} g_{\cdot i_2} g_{\cdot i_3} \left[S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 12}\right] + g_{\cdot i_1} g_{\cdot i_2} g_{\cdot i_3 i_4} \left[S_{i_1\cdot 1} S_{i_3\cdot 1}\right] \left[S_{i_2\cdot 1} S_{i_4\cdot 1}\right], \\
[111]_T &= g_{\cdot i}[111]_{S_i} + 3 g_{\cdot ij} \left[S_{i\cdot 1} S_{j\cdot 11}\right] + g_{\cdot ijk} \sum^{3} \left[S_{i\cdot 1} S_{j\cdot 1} S_{k\cdot 1}\right], \\
[1122]_T &= g_{\cdot i}[1122]_{S_i} + g_{\cdot i_1 i_2} \left(4 \left[S_{i_1\cdot 1} S_{i_2\cdot 122}\right] + 2 \left[S_{i_1\cdot 12} S_{i_2\cdot 12}\right] + [11]_{S_i} [11]_{S_j}\right) \\
&\quad + 2 g_{\cdot i_1 i_2 i_3} \left(2 \left[S_{i_1\cdot 1} S_{i_2\cdot 2} S_{i_3\cdot 12}\right] + \left[S_{i_1\cdot 1} S_{i_2\cdot 1}\right] [11]_{S_{i_3}}\right) \\
&\quad + g_{\cdot i_1\cdots i_4} \left[S_{i_1\cdot 1} S_{i_2\cdot 1}\right] \left[S_{i_3\cdot 1} S_{i_4\cdot 1}\right], \\
[1,122]_T &= g_{\cdot i} g_{\cdot j} \left[S_{i\cdot 1} S_{j\cdot 122}\right] + g_{\cdot i} g_{\cdot jk} \left(\left[S_{i\cdot 1} S_{j\cdot 1}\right] [11]_{S_k} + 2 \left[S_{i\cdot 1} S_{j\cdot 2} S_{k\cdot 12}\right]\right) \\
&\quad + g_{\cdot i} g_{\cdot jkl} \left[S_{i\cdot 1} S_{j\cdot 1}\right] \left[S_{k\cdot 1} S_{l\cdot 1}\right], \\
\left[12^2\right]_T &= g_{\cdot i} g_{\cdot j} \left[S_{i\cdot 12} S_{j\cdot 12}\right] + 2 g_{\cdot i} g_{\cdot jk} \left[S_{i\cdot 12} S_{j\cdot 1} S_{k\cdot 2}\right] + g_{\cdot ij} g_{\cdot kl} \left[S_{i\cdot 1} S_{k\cdot 1}\right] \left[S_{j\cdot 1} S_{l\cdot 1}\right],
\end{aligned}$$

and so on, where

$$[S_{i\cdot 1}S_{j\cdot 1}] = \int_{-\infty}^{\infty} S_{i\cdot 1}S_{j\cdot 1}dF\left(x_1\right),$$
$$[S_{i\cdot 1}S_{j\cdot 1}S_{k\cdot 12}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_{i\cdot 1}S_{j\cdot 1}S_{k\cdot 12}dF\left(x_1\right)dF\left(x_2\right),$$

and so on.

3 Some applications

Let $\widehat{F}$ be the empirical distribution of a random sample of size n from F. By Withers (1983), for a broad class of T, the cumulants of $T(\widehat{F})$ satisfy

$$\kappa_r\left(T\left(\widehat{F}\right)\right) \approx \sum_{i=r-1}^{\infty} n^{-i}a_{ri}$$

for $r \geq 1$, where the *cumulant coefficient* $a_{ri}(T) = a_{ri}$ is a certain function of the derivatives of $T(F)$. The most important are $a_{10} = T(F)$,

$$a_{21} = \left[1^2\right]_T = \left[T_{\cdot 1}^2\right], \tag{3.1}$$
$$a_{11} = [11]_T/2 = \left[T_{\cdot 11}\right]/2, \tag{3.2}$$
$$a_{32} = \left[1^3\right]_T + 3[1,2,12]_T = \left[T_{\cdot 1}^3\right] + 3\left[T_{\cdot 1}T_{\cdot 2}T_{\cdot 12}\right], \tag{3.3}$$

where

$$[f\left(T_{\cdot 1}, T_{\cdot 11}, \ldots\right)] = \int_{-\infty}^{\infty} f\left(T_{\cdot 1}, T_{\cdot 11}, \ldots\right)dF_1\left(x_1\right),$$
$$[f\left(T_{\cdot 1}, T_{\cdot 2}, T_{\cdot 11}, T_{\cdot 12}, T_{\cdot 22}, T_{\cdot 122}, \ldots\right)]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left(T_{\cdot 1}, T_{\cdot 2}, T_{\cdot 11}, T_{\cdot 12}, T_{\cdot 22}, T_{\cdot 122}, \ldots\right)dF_1\left(x_1\right)dF_2\left(x_2\right),$$

and so on, for $F_i = F(x_i)$, and

$$\left[1^i, 2^j, 11^k, 12^l, 22^m, \ldots\right]_T = \left[T_{\cdot 1}^i, T_{\cdot 2}^j, T_{\cdot 11}^k, T_{\cdot 12}^l, T_{\cdot 22}^m, \ldots\right],$$

and so on. We refer to the functionals $[\cdots]$ as *bracket functions*. They are the building blocks for the cumulant coefficients a_{ri} and the cumulant coefficients of the Studentized statistics, and hence for the Edgeworth-Cornish-Fisher expansions of the standardized form of $T(\widehat{F})$,

$$Y_n = \left(n/a_{21}\right)^{1/2}\left\{T\left(\widehat{F}\right) - T(F)\right\}, \tag{3.4}$$

and its Studentized form. They are also the building blocks for obtaining nonparametric confidence intervals and estimates of low bias for $T(F)$.

As a start we have these approximations to the bias, variance, and skewness of $T(\widehat{F})$:

$$\mathbb{E}\left[T\left(\widehat{F}\right)\right] = T(F) + n^{-1}a_{11} + O\left(n^{-2}\right),$$
$$\mathrm{var}\left[T\left(\widehat{F}\right)\right] = n^{-1}a_{21} + O\left(n^{-2}\right),$$
$$\mu_3\left[T\left(\widehat{F}\right)\right] = n^{-2}a_{32} + O\left(n^{-3}\right),$$

where $\mu_3[X] = \mathbb{E}[(X - \mathbb{E}[X])^3]$.

Theorem 3.1 lists the bracket functions needed for bias and bias reduction. Theorem 3.2 lists the bracket functions needed for Edgeworth-Cornish-Fisher expansions. Theorem 3.3 lists the bracket functions needed for nonparametric confidence intervals.

Theorem 3.1 *Under regularity conditions,*

$$\begin{aligned} \mathbb{E}\left[T\left(\widehat{F}\right)\right] &= T(F) + \sum_{i=1}^{j} n^{-i} a_{1i} + O\left(n^{-j-1}\right), \\ a_{11} &= [11]_T/2, \\ a_{12} &= [111]_T/6 + [1122]_T/8, \\ a_{13} &= [1111]_T/24 + [1122]_T/12 + [112233]_T/48, \end{aligned}$$

and so on. The estimates of $T(F)$ *of bias* $O(n^{-j-1})$ *are*

$$T\left(\widehat{F}\right) + \sum_{i=1}^{j} n^{-i} T_i\left(\widehat{F}\right) \text{ and } T\left(\widehat{F}\right) + \sum_{i=1}^{j} S_i\left(\widehat{F}\right)/(n-1)_i, \tag{3.5}$$

where $(m)_i = m!/(m-i)! = m(m-1)\cdots(m-i+1)$ *and*

$$T_1(F) = S_1(F) = -[11]_T/2, \tag{3.6}$$

$$T_2(F) = [111]_T/3 + [1122]_T/8 - [11]_T/2, \tag{3.7}$$

$$\begin{aligned} T_3(F) &= -[11]_T/2 + [111]_T - [1111]_T/4 + 3[1122]_T/4 - [11122]_T/6 - [112233]_T/48, \\ S_2(F) &= [111]_T/3 + [1122]_T/8, \\ S_3(F) &= -[1111]_T/4 + 3[1122]_T/8 - [11122]_T/6 - [112233]_T/48. \end{aligned}$$

Proof: Follows by equation (2.4) of Withers (1987). □

Theorem 3.2 *The 'reduced' Edgeworth and Cornish-Fisher expansions of* $T(\widehat{F})$ *to* $O(n^{-(j+1)/2})$ *needs*

$$\begin{aligned} &\text{for } j = 0 : a_{21}, \\ &\text{for } j = 1 : a_{11} \text{ and } a_{32}, \\ &\text{for } j = 2 : a_{22} = [1,11]_T + \left[12^2\right]_T/2 + [1,122]_T, \end{aligned}$$

$$a_{43} = \left[1^4\right]_T - 3\left[1^2\right]_T^2 + 12\left[1,2^2,12\right]_T + 12[1,2,13,23]_T + 4[1,2,3,123]_T, \tag{3.8}$$

and so on. In particular, for Y_n *of (3.4), under regularity conditions,*

$$P\left(Y_n \le x\right) = \Phi(x) - \phi(x)\left[n^{-1/2}h_1(x) + n^{-1}h_2(x)\right] + O\left(n^{-3/2}\right)$$

for $h_1 = A_{11} + A_{32}He_2/6$ *and* $h_1 = (A_{22} + A_{11}^2)He_1/2 + (A_{43} + 4A_{11}A_{32})He_3/24 + A_{32}^2 He_5/72$, *where* Φ, ϕ *are the distribution and density of a unit normal random variable,* He_r *is the* r*th Hermite polynomial, and* $A_{ri} = a_{ri}/a_{21}^{r/2}$, *the standardized cumulant coefficient.*

Proof: Follows by Withers (1983). □

The regularity conditions needed for Theorems 3.1 and 3.2 are the same as those given in Withers (1983, 1987). So, they are not stated here.

Theorem 3.3 *A confidence interval for $T(F)$ of level $1-\alpha+O(n^{-(j+1)/2})$ requires the bracket functions*

$$\begin{aligned}
&\textit{for } j=0: a_{21}=\left[1^2\right]_T,\\
&\textit{for } j=1: [11]_T,\ \left[1^3\right]_T,\ [1,2,12]_T,\\
&\textit{for } j=2: [1,11]_T, \left[12^2\right]_T, [1,122]_T, \left[1^4\right]_T, \left[1,2^2,12\right]_T,\\
&\qquad\qquad [1,2,13,23]_T, [1,2,3,123]_T.
\end{aligned} \tag{3.9}$$

Proof: Follows by Theorem 5.1 in Withers (1983). $\square$

By Withers (1989), the bracket functions in Theorem 3.3 are also the terms needed for the distribution and quantiles of the Studentized form of $T(F)$ to $O(n^{-(j+1)/2})$.

For the distribution of $\mid T(\widehat{F})-T(F)\mid$ to $O(n^{-j-1})$ or for a symmetric confidence interval for $T(F)$ of level $1-\alpha+O(n^{-j-1})$ one needs, by equations (2.4) and (2.5) of Withers (1982), a_{21} for $j=0$ and a_{11}, a_{32}, a_{22}, a_{43} for $j=1$.

For convenience, set $T=T(F)$ and $g_i=g^{(i)}(S(F))$ for $S(F)$ in $\mathbb{R}$.

Example 3.1 *This example gives bracket functions for a function of a univariate functional. Suppose (2.6) holds with $b=1$. Then*

$$\begin{aligned}
T_{\cdot 1} &= g_1 S_{\cdot 1},\\
T_{\cdot 12} &= g_1 S_{\cdot 12}+g_2 S_{\cdot 1}S_{\cdot 2},\\
T_{\cdot 123} &= g_1 S_{\cdot 123}+g_2\sum^{3} S_{\cdot 1}S_{\cdot 23}+g_3 S_{\cdot 1}S_{\cdot 2}S_{\cdot 3},\\
T_{\cdot 1234} &= g_1 S_{\cdot 1234}+g_2\left(\sum^{4} S_{\cdot 1}S_{\cdot 234}+\sum^{3} S_{\cdot 12}S_{\cdot 34}\right)\\
&\quad +g_3\sum^{6} S_{\cdot 1}S_{\cdot 2}S_{\cdot 34}+g_4 S_{\cdot 1}S_{\cdot 2}S_{\cdot 3}S_{\cdot 4},\\
T_{\cdot 12345} &= g_1 S_{\cdot 1\cdots 5}+g_2\left(\sum^{4} S_{\cdot 1}S_{\cdot 2345}+\sum^{10} S_{\cdot 12}S_{\cdot 345}\right)\\
&\quad +g_3\left(\sum^{10} S_{\cdot 1}S_{\cdot 2}S_{\cdot 345}+\sum^{15} S_{\cdot 1}S_{\cdot 23}S_{\cdot 45}\right)+g_4\sum^{10} S_{\cdot 1}S_{\cdot 2}S_{\cdot 3}S_{\cdot 45}+g_5 S_{\cdot 1}\cdots S_{\cdot 5}.
\end{aligned}$$

So,

$$\begin{aligned}
&\left[1^k\right]_T = g_1^k\left[1^k\right]_S,\\
&[11]_T = g_1[11]_S+g_2\left[1^2\right]_S,\\
&[1,2,12]_T = g_1^3[1,2,12]_S+g_1^2 g_2\left[1^2\right]_S^2,\\
&[111]_T = g_1[111]_S+3g_2[1,11]_S+g_3\left[1^3\right]_S,\\
&[1122]_T = g_1[1122]_S+g_2\left(4[1,122]_S+[11]_S^2+2\left[12^2\right]_S\right)\\
&\qquad +2g_3\left(\left[1^2\right]_S[11]_S+2[1,2,12]_S\right)+g_4\left[1^2\right]_S^2,\\
&[1,122]_T = g_1^2[1,122]_S+g_1 g_2\left(\left[1^2\right]_S[111]_S+2[1,2,12]_S\right)+g_1 g_3\left[1^2\right]_S^2,\\
&\left[12^2\right]_T = g_1^2\left[12^2\right]_S+2g_1 g_2[1,2,12]_S+g_2^2\left[1^2\right]_S^2.
\end{aligned}$$

Example 3.2 *This example gives bracket functions for a product. Suppose that $T(F) = S_1(F)S_2(F)$. Then*

$$\begin{aligned}
T_{\cdot 1} &= S_2 S_{1\cdot 1} + S_1 S_{2\cdot 1} = \sum^{(2)} S_1 S_{1\cdot 1} \text{ say,} \\
T_{\cdot 12} &= (S_2 S_{1\cdot 12} + S_1 S_{2\cdot 12}) + (S_{1\cdot 1} S_{2\cdot 2} + S_{2\cdot 1} S_{1\cdot 2}) \\
&= \sum^{(2)} (S_2 S_{1\cdot 12} + S_1 S_{2\cdot 12}) \text{ say,} \\
T_{\cdot 123} &= (S_2 S_{1\cdot 123} + S_1 S_{2\cdot 123}) + \sum^{3} (S_{1\cdot 1} S_{2\cdot 23} + S_{2\cdot 1} S_{1\cdot 23}) \\
&= \sum^{(2)} S_2 S_{1\cdot 123} + \sum^{(6)} S_{1\cdot 1} S_{2\cdot 23} \text{ say,} \\
T_{\cdot 1234} &= (S_2 S_{1\cdot 1234} + S_1 S_{2\cdot 1234}) + \sum^{4} (S_{1\cdot 1} S_{2\cdot 234} + S_{2\cdot 1} S_{1\cdot 234}) \\
&\quad + \sum^{3} (S_{1\cdot 12} S_{2\cdot 34} + S_{2\cdot 12} S_{1\cdot 34}) \\
&= \sum^{(2)} S_2 S_{1\cdot 1234} + \sum^{(8)} S_{1\cdot 1} S_{2\cdot 234} + \sum^{6} S_{1\cdot 2} S_{2\cdot 34} \text{ say.}
\end{aligned}$$

So,

$$\begin{aligned}
\left[1^2\right]_T &= \sum^{(2)} S_2^2 \left[1^2\right]_{S_1} + 2 S_1 S_2 \left[S_{1\cdot 1} S_{2\cdot 1}\right], \\
[11]_T &= \sum^{(2)} S_2 [11]_{S_1} + 2 \left[S_{1\cdot 1} S_{2\cdot 1}\right], \\
\left[1^3\right]_T &= \sum^{(2)} \left(S_2^3 \left[1^3\right]_{S_1} + 3 S_2^2 \left[S_{2\cdot 1} S_{1\cdot 1}^2\right] \right), \\
[1,2,12]_T &= \sum^{(2)} \Bigg\{ S_2^3 [1,2,12]_{S_1} + S_2^2 S_1 \Bigg(\left[S_{1\cdot 1} S_{1\cdot 2} S_{2\cdot 12}\right] \\
&\quad + 2 \left[S_{1\cdot 1} S_{2\cdot 2} S_{1\cdot 12}\right] \Bigg) + 2 S_2^2 \left[1^2\right]_{S_1} \left[S_{1\cdot 1} S_{2\cdot 1}\right] \Bigg\} \\
&\quad + 2 S_1 S_2 \left(\left[S_{1\cdot 1} S_{2\cdot 1}\right]^2 + \left[1^2\right]_{S_1} \left[1^2\right]_{S_2} \right), \\
[111]_T &= \sum^{(2)} \left(S_2 [111]_{S_1} + 3 \left[S_{2\cdot 1} S_{1\cdot 11}\right] \right), \\
[1122]_T &= \sum^{(2)} \left(S_2 [1122]_{S_1} + 4 \left[S_{1\cdot 1} S_{2\cdot 122}\right] \right) + 2 [11]_{S_1} [11]_{S_2} + 4 \left[S_{1\cdot 12} S_{2\cdot 12}\right].
\end{aligned}$$

Example 3.3 *This example gives bracket functions for a Studentized function. The Studentized form of $T(F)$ is*

$$T_0\left(\widehat{F}\right) = V\left(\widehat{F}\right)^{-1/2} \left\{ T\left(\widehat{F}\right) - T(F) \right\}$$

for $V(F) = a_{21}$. Its bracket functions $[\cdots]_{T_0}$ (and so also its cumulant coefficients) may be expressed in terms of the bracket functions $[\cdots]_T$. For details, see Appendix A of Withers (1989).

If one makes other assumptions such as symmetry of F or a parametric form for F, then $V(F) = a_{21}$ will generally take a simpler form. Similarly, in some circumstances one is interested in standardizing a functional in a different way, for example, replacing μ_r by $\mu_r/\mu_2^{r/2}$. The next example covers this situation for the special case of a $T(F)$ a function of a univariate functional.

Example 3.4 *Suppose that $T(F) = U(F)g(S(F))$ with $S(F)$ in $\mathbb{R}$. Then*

$$
\begin{aligned}
T_{\cdot 1} &= g_0 U_{\cdot 1} + g_1 U S_{\cdot 1}, \\
T_{\cdot 12} &= g_0 U_{\cdot 12} + g_1 \left(U S_{\cdot 12} + \sum^{2} U_{\cdot 1} S_{\cdot 2} \right) + g_2 U S_{\cdot 1} S_{\cdot 2}, \\
T_{\cdot 123} &= g_0 U_{\cdot 123} + g_1 \left(U S_{\cdot 123} + \sum^{(6)} U_{\cdot 1} S_{\cdot 23} \right) \\
&\quad + g_2 \sum^{3} \left(U S_{\cdot 1} S_{\cdot 23} + U_{\cdot 1} S_{\cdot 2} S_{\cdot 3} \right) + g_3 U S_{\cdot 1} S_{\cdot 2} S_{\cdot 3}, \\
T_{\cdot 1234} &= g_0 U_{\cdot 1234} + g_1 \left(U S_{\cdot 1234} + \sum^{(8)} U_{\cdot 1} S_{\cdot 234} + \sum^{6} U_{\cdot 12} S_{\cdot 34} \right) \\
&\quad + g_2 \left(U \sum^{4} S_{\cdot 1} S_{\cdot 234} + U \sum^{3} S_{\cdot 12} S_{\cdot 34} + \sum^{12} U_{\cdot 1} S_{\cdot 2} S_{\cdot 34} + \sum^{6} U_{\cdot 12} S_{\cdot 3} S_{\cdot 4} \right) \\
&\quad + g_3 \left(U \sum^{6} S_{\cdot 1} S_{\cdot 2} S_{\cdot 34} + \sum^{4} U_{\cdot 1} S_{\cdot 2} S_{\cdot 3} S_{\cdot 4} \right) + g_4 U S_{\cdot 1} S_{\cdot 2} S_{\cdot 3} S_{\cdot 4}.
\end{aligned}
$$

So, the cumulant coefficients a_{21}, a_{11}, a_{32} needed for third order inference are given by (3.1)-(3.3) in terms of the bracket functions

$$
\begin{aligned}
\left[1^2\right]_T &= g_0^2 \left[1^2\right]_U + 2 g_0 g_1 U \left[U_{\cdot 1} S_{\cdot 1}\right] + g_1^2 U^2 \left[1^2\right]_S, \\
[11]_T &= g_0 [11]_U + g_1 \left(U [11]_S + 2 \left[U_{\cdot 1} S_{\cdot 1}\right] \right) + g_2 U \left[1^2\right]_S, \\
\left[1^3\right]_T &= g_0^3 \left[1^3\right]_U + 3 g_0^2 g_1 U \left[U_{\cdot 1}^2 S_{\cdot 1}\right] + 3 g_0 g_1^2 U^2 \left[U_{\cdot 1} S_{\cdot 1}^2\right] + g_1^3 U^3 \left[1^3\right]_S, \\
[1,2,12]_T &= g_0^3 [1,2,12]_U + g_0^2 g_1 \left(U \left[U_{\cdot 1} U_{\cdot 2} S_{\cdot 12}\right] + 2 \left[1^2\right]_U \left[U_{\cdot 1} S_{\cdot 1}\right] + 2U \left[U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}\right] \right) \\
&\quad + g_0^2 g_2 U \left[U_{\cdot 1} S_{\cdot 1}\right]^2 + g_0 g_1^2 U \Big(2 \left[U_{\cdot 1} S_{\cdot 1}\right]^2 + 2 \left[1^2\right]_U \left[1^2\right]_S + 2U \left[U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}\right] \\
&\quad + U \left[S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}\right] \Big) + 2 g_0 g_1 g_2 U^2 \left[U_{\cdot 1} S_{\cdot 1}\right] \left[1^2\right]_S + g_1^3 U^2 \Big(U [1,2,12]_S \\
&\quad + 2 \left[U_{\cdot 1} S_{\cdot 1}\right] \left[1^2\right]_S \Big) + g_1^2 g_2 U^3 \left[1^2\right]_S^2.
\end{aligned}
$$

Similarly, the cumulant coefficients a_{22}, a_{43} needed for third order inference are given by (3.8) in terms of the bracket functions given in Appendix A. The bracket functions needed

for (3.6), (3.7) for estimates of $T(F)$ of bias $O(n^{-3})$ are

$$\begin{aligned}
[111]_T &= g_0\,[111]_U + g_1\left(U\,[111]_S + 3\,[U_{\cdot 1}S_{\cdot 11}] + 3\,[U_{\cdot 11}S_{\cdot 1}]\right) \\
&\quad +3g_2\left(U\,[1,11]_S + \left[U_{\cdot 1}S_{\cdot 1}^2\right]\right) + g_3 U\left[1^3\right]_S, \\
[1122]_T &= g_0\,[1122]_U + g_1\Bigg(U\,[1122]_S + 4\,[U_{\cdot 1}S_{\cdot 122}] + 4\,[U_{\cdot 122}S_{\cdot 1}] \\
&\quad +2\,[11]_U\,[11]_S + 4\,[U_{\cdot 12}S_{\cdot 12}]\Bigg) + g_2\Bigg(4U\,[1,122]_S + U\,[11]_S^2 + 2U\left[12^2\right]_S \\
&\quad +8\,[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] + 4\,[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] + 4\,[U_{\cdot 1}S_{\cdot 1}]\,[11]_S + 2\,[11]_U\left[1^2\right]_S\Bigg) \\
&\quad +2g_3\left(U\left[1^2\right]_S[11]_S + 2U\,[1,2,12]_S + 2\,[U_{\cdot 1}S_{\cdot 1}]\left[1^2\right]_S\right) + g_4 U\left[1^2\right]_S^2.
\end{aligned}$$

Further terms are given in Appendix A.

If $g(s) = s^r$ then $g_i = (r)_i S^{r-i}$. Putting $r = -1$ gives the derivatives of a quotient $(-1)_i = (-1)^i i!$.

4 Applications to moments

Suppose $X \sim F$ on $\mathbb{R}$. Set $\mu = \mathbb{E}[X]$, $\mu_r^{'} = \mathbb{E}[X^r]$ and let $\{\mu_r, \kappa_r\}$ be the central moments and cumulants of F. Set $\mu(F) = \mu$ and so on. Let $\widehat{F}$ be the empirical distribution of a random sample of size n from F.

Many authors have studied problems of moments and cumulants: see, for example, Stuart and Ord (1987). Fisher's *k-statistic* k_r, the unbiased estimate of κ_r, is given there in Section 12.9 for $r \leq 8$ in terms of $\{s_i = n\mu_i^{'}(\widehat{F}) = \sum_{j=1}^n X_j^i\}$. Fisher's expressions for unbiased estimates of the joint cumulants of k-statistics are given there in Section 12.16. Wishart's unbiased estimates of products of cumulants are given there in Section 12.16 in terms of symmetric functions, which can be converted to $\{s_i\}$ using Appendix Table 10.

Generally one only wants approximations. (Indeed without making parametric assumptions on F only approximations are possible except for estimating polynomials in moments). One problem with these "traditional" approaches is that it is not an easy task to separate out terms beyond the first in decreasing order of importance in order to make such approximations. As noted in Section 3 the present approach does not suffer from this disadvantage.

For $S(F)$ a polynomial in F of degree r (for example, $\mu_r^{'}$, μ_r or κ_r), derivatives of order beyond r vanish.

Example 4.1 *Suppose $T(F)$ is a function of a univariate mean, say $T(F) = g(\mu(F))$. Setting $g_k = g^{(k)}(\mu)$, Example 3.1 implies*

$$\begin{aligned}
a_{21} &= g_1^2\mu_2, \\
a_{11} &= g_2\mu_2/2, \\
a_{32} &= g_1^3\mu_3 + 3g_1^2 g_2\mu_2^2, \\
a_{22} &= g_1 g_2\mu_3 + \left(g_2^2/2 + g_1 g_3\right)\mu_2^2, \\
a_{43} &= g_1^4\left(\mu_4 - 3\mu_2^2\right) + 12g_1^3 g_2\mu_3\mu_2 + 4\left(3g_1^2 g_2^2 + g_1^3 g_3\right)\mu_2^3.
\end{aligned}$$

For,

$$T_{\cdot 1\cdots p} = g_p h_1 \cdots h_p,$$

where $h_i = x_i - \mu$. *So,*

$$\begin{aligned}
\left[1^k\right] &= g_1^k \mu_k, \\
[1\cdots 1] &= g_k \mu_k \textit{ if } 1\cdots 1 \textit{ contains } k \textit{ 1's}, \\
[1,2,12] &= g_1^2 g_2 \mu_2^2, \\
[1,11] &= g_1 g_2 \mu_3, \\
\left[12^2\right] &= g_2^2 \mu_2^2, \\
[1,122] &= g_1 g_3 \mu_2^2, \\
\left[1,2^2,12\right] &= g_1^3 g_2 \mu_3 \mu_2, \\
[1,2,13,23] &= g_1^2 g_2^2 \mu_2^3, \\
[1,2,3,123] &= g_1^3 g_3 \mu_2^3.
\end{aligned}$$

So, an estimate of $g(\mu(F))$ *of bias* $O(n^{-4})$ *is given by (3.5) with* $j=3$ *in terms of*

$$\begin{aligned}
S_{\cdot 1}(F) &= -g_2\mu_2/2, \\
S_{\cdot 2}(F) &= g_3\mu_3/3 + g_4\mu_2^2/8, \\
S_{\cdot 3}(F) &= -g_4\mu_4/4 + 3g_4\mu_2^2/8 - g_5\mu_3\mu_2/6 - g_6\mu_2^3/48.
\end{aligned}$$

For example, an estimate of μ^r *of bias* $O(n^{-4})$ *is given by substituting* $g_i = (r)_i \mu^{r-i}$. *If* $\mu \geq 0$, r *need not be an integer. However, regularity conditions generally breakdown if* $r < 0$ *and* $\dot{F}(0) \neq 0$.

Functions of non-central moments can be handled with similar ease. We now present an important result which was stated without proof in equation (4.1) of Withers (1987), *the derivatives of a central moment.*

Theorem 4.1 *For* r, p *in* $\{1,2,\ldots\}$, *the pth derivative of* $\mu_r(F)$ *is*

$$\mu_{r\cdot 1\cdots p} = (-1)^p \left\{ (r)_p \mu_{r-p} - (r)_{p-1} \sum_{i=1}^p \left(h_i^{r-p} - \mu_{r-p+1} h_i^{-1} \right) \right\} \prod_{i=1}^p h_i,$$

where $h_i = x_i - \mu$.

Proof: As in Example 3.1, $T(F) = \mu_k^{'} \mu^j$ has derivatives

$$T_{\cdot 1\cdots p} = \left(g_{p-1} \sum_{i=1}^p U_{\cdot i} S_{\cdot i}^{-1} + g_p U \right) S_{\cdot 1} \cdots S_{\cdot p},$$

where

$$g_p = (j)_p \mu^{j-p}, \ U = \mu_k^{'}, \ U_{\cdot i} = x_i^k - \mu_k^{'}, \ S_{\cdot i} = x_i - \mu.$$

But $\mu_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} \mu_k^{'} \mu^{r-k}$. So,

$$\begin{aligned}\mu_{r \cdot 1 \cdots p} &= \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} \left\{ (r-k)_{p-1} \mu^{r-k-p+1} \sum_{i=1}^{p} \left(x_i^k - \mu_k^{'} \right) h_i^{-1} \right. \\ &\qquad \left. + (r-k)_p \mu^{r-k-p} \mu_k^{'} \right\} \prod_{j=1}^{p} h_j.\end{aligned}$$

Now simplify. □

Some particular cases of the theorem are given by the following corollaries.

Corollary 4.1 *We have*

$$\begin{aligned}
&\mu_{r \cdot 1} = h_1^r - \mu_r - r h_1 \mu_{r-1}, \\
&\mu_{r \cdot 12} = -r \sum^{2} \left(h_1^{r-1} - \mu_{r-1} \right) h_2 + (r)_2 h_1 h_2 \mu_{r-2}, \\
&\mu_{r \cdot 123} = (r)_2 \sum^{3} \left(h_1^{r-2} - \mu_{r-2} \right) h_2 h_3 - (r)_3 h_1 h_2 h_3 \mu_{r-3}, \\
&\mu_{r \cdot 12 \cdots r-1} = (-1)^{r-1} (r!/2) \sum^{r-1} \left(h_1^2 - \mu_2 \right) h_2 \cdots h_{r-1}, \\
&\mu_{r \cdot 12 \cdots r} = (-1)^r (r-1) r! h_1 \cdots h_r, \\
&\mu_{r \cdot 1^p} = (-1)^p \left\{ (r)_p \mu_{r-p} h_1^p - p (r)_{p-1} h_1^r - \mu_{r-p+1} h_1^{p-1} \right\}.
\end{aligned}$$

Corollary 4.2 *We have*

$$\begin{aligned}
&\mu_{2 \cdot 1} = h_1^2 - \mu_2, \\
&\mu_{3 \cdot 1} = h_1^3 - \mu_3 - 3 h_1 \mu_2, \\
&\mu_{4 \cdot 1} = h_1^4 - \mu_4 - 4 h_1 \mu_3, \\
&\mu_{5 \cdot 1} = h_1^5 - \mu_5 - 5 h_1 \mu_4, \\
&\mu_{2 \cdot 12} = -2 h_1 h_2, \\
&\mu_{3 \cdot 12} = -3 \sum^{2} \left(h_1^2 - \mu_2 \right) h_2, \\
&\mu_{4 \cdot 12} = 12 h_1 h_2 \mu_2 - 4 \sum^{2} \left(h_1^3 - \mu_3 \right) h_2, \\
&\mu_{5 \cdot 12} = 20 h_1 h_2 \mu_3 - 5 \sum^{2} \left(h_1^4 - \mu_4 \right) h_2, \\
&\mu_{3 \cdot 123} = 12 h_1 h_2 h_3, \\
&\mu_{4 \cdot 123} = 12 \sum^{3} \left(h_1^2 - \mu_2 \right) h_2 h_3, \\
&\mu_{5 \cdot 123} = 20 \sum^{3} \left(h_1^3 - \mu_3 \right) h_2 h_3 - 60 h_1 h_2 h_3, \\
&\mu_{4 \cdot 1234} = 72 h_1 h_2 h_3 h_4, \\
&\mu_{5 \cdot 1234} = 60 \sum^{4} \left(h_1^2 - \mu_2 \right) h_2 h_3 h_4, \\
&\mu_{5 \cdot 12345} = -480 h_1 h_2 h_3 h_4 h_5.
\end{aligned}$$

So, for example, for $T(F) = \mu_r$,

$$\begin{aligned}
\left[1^3\right]_T &= \mu_{3r} - 3r\mu_{2r+1}\mu_{r-1} - 3\mu_{2r}\mu_r + 3r^2\mu_{r+2}\mu_{r-1}^2 + 6r\mu_{r+1}\mu_r\mu_{r-1} \\
&\quad +2\mu_r^3 - 3r^2\mu_r\mu_{r-1}^2\mu_2 - r^3\mu_{r-1}^3\mu_3, \\
[1,2,12]_T &= -\mu_{2r-1}\left(2r\mu_{r+1} + r^2\mu_{r-1}\mu_2\right) + (r)_2\mu_{r+1}^2\mu_{r-2} \\
&\quad -r^2\mu_{r+1}\mu_r\mu_{r-1} + 2r^2(r-1)\mu_{r+1}\mu_{r-1}\mu_{r-2}\mu_2 \\
&\quad -\left(2r^3 - r^2\right)\mu_r\mu_{r-1}^2\mu_2 + r^3(r-1)\mu_{r-1}^2\mu_{r-2}\mu_2^2,
\end{aligned}$$

giving

$$\begin{aligned}
a_{21} &= \left[1^2\right]_T = r^2\mu_{r-1}^2\mu_2 - 2r\mu_{r-1}\mu_{r+1} + \mu_{2r} - \mu_r^2, \\
a_{11} &= [11]_T/2 = (r)_2\mu_{r-2}\mu_2/2 - r\mu_r, \\
a_{32} &= \mu_{3r} - 3r\mu_{2r+1}\mu_{r-1} - 3\mu_{2r}\mu_r - 3\mu_{2r-1}\left(2r\mu_{r+1} + r^2\mu_{r-1}\mu_2\right) \\
&\quad +3r^2\mu_{r+2}\mu_{r-1}^2 + 3(r)_2\mu_{r+1}^2\mu_{r-2} - 3r(r-2)\mu_{r+1}\mu_r\mu_{r-1} \\
&\quad +6r^2(r-1)\mu_{r+1}\mu_{r-1}\mu_{r-2}\mu_2 + 2\mu_r^3 - 3r^2\mu_r\mu_{r-1}^2\mu_2 - r^3\mu_{r-1}^3\mu_3.
\end{aligned}$$

Similarly, estimates of μ_r for *general* r of bias $O(n^{-3})$ are given by (3.5) in terms of

$$\begin{aligned}
&[111]_T = -(r)_3\mu_{r-3}\mu_3 + 3(r)_2\left(\mu_r - \mu_{r-2}\mu_2\right), \\
&[1122]_T = (r)_4\mu_{r-4}\mu_2^2.
\end{aligned}$$

For $r = 2$ this gives

$$a_{21} = \mu_4 - \mu_2^2,\ a_{11} = -\mu_2,\ a_{32} = \mu_6 - 3\mu_4\mu_2 + 2\mu_2^3,\ [111]_T = [1122]_T = 0,$$

and for $r = 3$ this gives

$$\begin{aligned}
&a_{21} = \mu_6 - 4\mu_4\mu_2 - \mu_3^2 + 9\mu_2^3,\ a_{11} = -3\mu_3, \\
&a_{32} = \mu_9 - 9\mu_7\mu_2 - 3\mu_6\mu_3 - 18\mu_5\mu_4 - 9\mu_4\mu_3\mu_2 + 2\mu_3^3, \\
&[111]_T = 12\mu_3,\ [1122]_T = 0.
\end{aligned}$$

Example 4.2 *Suppose that r is an odd integer and F is symmetric. So, odd cumulants of $\mu_r(\widehat{F})$ are zero so that $a_{11} = a_{32} = 0$ and the Edgeworth-Cornish-Fisher expansions are in powers of n^{-1}, not just $n^{-1/2}$. Taking $r = 3$ gives for $T(F) = \mu_3$,*

$$\begin{aligned}
&\left[12^2\right]_T = 2\mu_2\left(\mu_4 - \mu_2^2\right), \\
&[1,11]_T = -6\left(\mu_6 - 4\mu_4\mu_2 + 3\mu_2^3\right), \\
&[1,122]_T = 12\mu_2\left(\mu_4 - 3\mu_2^2\right), \\
&\left[1^4\right]_T = \mu_{12} - 12\mu_{10}\mu_2 + 5\mu_8\mu_2^2 - 108\mu_6\mu_2^3 + 81\mu_4\mu_2^4, \\
&\left[1,2^2,12\right]_T = -3\mu_4\left(\mu_8 - 4\mu_6\mu_2 + 6\mu_4\mu_2^2 - 3\mu_2^4\right), \\
&[1,2,13,23]_T = 9\left(\mu_4 - \mu_2^2\right)\left(\mu_4 - 3\mu_2^2\right)^2.
\end{aligned}$$

So,

$$\begin{aligned}
&a_{21} = \mu_6 - 6\mu_4\mu_2 + 9\mu_2^3, \\
&a_{22} = -6\left(\mu_6 - 7\mu_4\mu_2 + 10\mu_2^3\right), \\
&a_{43} = \mu_{12} - 12\mu_{10}\mu_2 - \mu_8\left(72\mu_4 - 5\mu_2^2\right) - 3\mu_6\left(\mu_6 - 108\mu_4\mu_2 + 54\mu_2^3\right) \\
&\qquad +3\mu_4\left(52\mu_4^2 - 576\mu_4\mu_2^2 + 1179\mu_2^4\right) - 2511\mu_2^6.
\end{aligned}$$

For F normal this gives $a_{21} = 6\mu_2^3$, $a_{22} = -24\mu_2^3$, $a_{43} = -11625\mu_2^6$.

Example 4.3 *This example is about standardized central moments. Suppose $T(F) = \nu_r$, where $\nu_r = \mu_r/\mu_2^{r/2}$. Then the $[\cdot]_T$ needed for third order inference and bias reduction, are given by Example 3.4 with $S = \mu_2$ and $U = \mu_r$ and*

$$g_j = (-r/2)_j \mu_2^{-r/2-j} = r(r+2)(r+4)\cdots(r+2j-2)\,(-2\mu_2)^{-j}\,\mu_2^{-r/2}$$

in terms of $[1^2]_U$, $[11]_U$, $[1^3]_U$, $\cdots$ and $[1^2]_S$, $[11]_S$, $[1^3]_S$, $\cdots$ given by Example 4.2 and the bracket functions

$$\begin{aligned}
&[U_{\cdot 1}S_{\cdot 1}] = \mu_{r+2} - 2\mu_r\mu_2 - r\mu_{r-1}\mu_3,\\
&\left[U_{\cdot 1}^2 S_{\cdot 1}\right] = \mu_{2r+2} - \mu_{2r}\mu_2 - 2r\mu_{r-1}\left(\mu_{r+3} - \mu_{r+1}\mu_2 - \mu_r\mu_3\right) - 2\mu_{r+2}\mu_r\\
&\qquad +2\mu_r^2\mu_2 + r^2\mu_{r-1}^2\left(\mu_4 - \mu_2^2\right),\\
&\left[U_{\cdot 1}S_{\cdot 1}^2\right] = \mu_{r+4} - 2\mu_{r+2}\mu_2 + \mu_r\left(-\mu_4 + 2\mu_2^2\right) - r\mu_{r-1}\left(\mu_5 + 2\mu_3\mu_2\right),\\
&[U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}] = -2\left(\mu_{r+1} - r\mu_{r-1}\mu_2\right)^2,\\
&[U_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = -r\mu_{2r-1}\mu_3 - r\mu_{r+1}^2 + \left(r^2+r\right)\mu_{r+1}\mu_{r-1}\mu_2 + (r)_2\,\mu_{r+1}\mu_{r-2}\mu_3\\
&\qquad + \left(r^2+r\right)\mu_r\mu_{r-1}\mu_3 - r^2\mu_{r-1}^2\mu_2^2 - r^2\,(r-1)\,\mu_{r-1}\mu_{r-2}\mu_3\mu_2,\\
&[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] = 2\left(-\mu_{r+1}\mu_3 + r\mu_{r-1}\mu_3\mu_2\right),\\
&[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = -2r\mu_{r+1}\mu_3 + 4r\mu_{r-1}\mu_3\mu_2 + (r)_2\,\mu_{r-2}\mu_3^2,\\
&[U_{\cdot 1}S_{\cdot 11}] = 2\left(-\mu_{r+2} + \mu_r\mu_2 + r\mu_{r-1}\mu_3\right),\\
&[U_{\cdot 11}S_{\cdot 1}] = 2r\left(-\mu_{r+2} + \mu_r\mu_2 + \mu_{r-1}\mu_3\right) + (r)_2\,\mu_{r-2}\left(\mu_4 - \mu_2^2\right),\\
&[U_{\cdot 1}S_{\cdot 122}] = 0,\\
&[U_{\cdot 122}S_{\cdot 1}] = (r)_2\left(\mu_r\mu_2 + 2\mu_{r-1}\mu_3 - \mu_{r-2}\mu_2^2 - (r-2)\,\mu_{r-3}\mu_3\mu_2\right),\\
&[U_{\cdot 12}S_{\cdot 12}] = 4r\mu_r\mu_2 - 2\,(r)_2\,\mu_{r-2}\mu_2^2.
\end{aligned}$$

For example, suppose that $r = 3$ and F is symmetric. Then $a_{ri} = 0$ for r odd and

$$\begin{aligned}
a_{21} &= \nu_6 - 6\nu_4 + 9,\\
a_{22} &= -3\left(\nu_8 - 5\nu_6 + 7\nu_4 - 3\right) + 12\nu_6\left(2\nu_4 - 1\right)/4\\
&\quad +2\nu_4\left(107\nu_4 - 489\right)/4 + 9\left(4\nu_4 - 11\right),\\
a_{43} &= \nu_{12} - 12\nu_{10} + 54\nu_8 - 108\nu_6 + 81\nu_4 - 3a_{21}^2\\
&\quad -18\left(\nu_8 - 4\nu_6 + 3\nu_4 - 3\right)\left(\nu_6 - 4\nu_4 + 9\right)\\
&\quad +27\left(\nu_4 - 1\right)\left[a_{21}^2 + 4a_{21}\left(\nu_4 - 3\right) + 4\left(\nu_6\left(\nu_4 - 1\right) + 9\right)\right]\\
&\quad +12\left(\nu_4 - 3\right)^2\left(3\nu_6 - 14\nu_4 + 15\right).
\end{aligned}$$

For F normal this is in agreement with Fisher (1931) who gave the result

$$\mu\left(-r, a_3, a_4, \ldots\right) = \mu\left(a_3, a_4, \ldots\right)(n-1)^r / \left\{(n-1)(n+1)\cdots(n+2r-3)\mu_2^r\right\}$$

for $\mu(a_2, a_3, a_4, \ldots) = \mathbb{E}[k_2^{a_2}k_3^{a_3}\cdots]$. See Agostino and Pearson (1973) for a simulation approach.

Figure 4.1 compares the bias reduced estimator of ν_3 versus the usual one by means of simulation. The biases of the estimators are computed by simulating ten thousand replications of samples of size n from the following distributions: standard normal, Student's t with two degrees of freedom, Student's t with five degrees of freedom, Student's t with ten

degrees of freedom, standard logistic, standard Laplace. As expected, the bias reduced estimators give substantially smaller biases for each n and for each of the six distributions. The biases appear largest for the Student's t distribution with two degrees of freedom. The biases appear smallest for the normal distribution, the Student's t distribution with ten degrees of freedom, and the logistic distribution.

As noted k_r is the unbiased estimate of κ_r so $k_2 = \mu_2(\widehat{F})n/(n-1)$, and $k_3 = \mu_3(\widehat{F})n^2/(n-1)(n-2)$.

Example 4.4 *Suppose $T(F) = \mu_r/\mu^r$, where $\mu \neq 0$. Then the $[\cdot]_T$ needed for third order inference and bias reduction are given by Example 3.4 with $g_j = (-r)_j\mu^{-r-j}$, $S = \mu$, $U = \mu_r$, $[1^2]_U$, $[11]_U$, $[1^3]_U$, $\ldots$ given by Example 4.2, $[1^i]_S = \mu_i$, the other non-zero leading terms needed for Example 3.3 being*

$$\begin{aligned}
&[U_{\cdot 1}S_{\cdot 1}] = \mu_{r+1} - r\mu_{r-1}\mu_2,\\
&\left[U_{\cdot 1}^2 S_{\cdot 1}\right] = \mu_{2r+1} - 2r\mu_{r+2}\mu_{r-1} - 2\mu_{r+1}\mu_r + 2r\mu_r\mu_{r-1}\mu_2 + r^2\mu_{r-1}^2\mu_3,\\
&\left[U_{\cdot 1}S_{\cdot 1}^2\right] = \mu_{r+2} - \mu_r\mu_2 - r\mu_{r-1}\mu_3,\\
&[U_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = -r\mu_{2r-1}\mu_2 - r\mu_{r+1}\mu_r + r(2r+1)\mu_r\mu_{r-1}\mu_2\\
&\qquad +(r)_2\mu_{r-2}\left(\mu_{r+1}\mu_2 - r\mu_{r-1}\mu_2^2\right),\\
&[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = -2r\mu_r\mu_2 + (r)_2\mu_{r-2}\mu_2^2,\\
&[U_{\cdot 11}S_{\cdot 1}] = -2r\left(\mu_{r+1} - \mu_{r-1}\mu_2\right) + (r)_2\mu_{r-2}\mu_3,\\
&[U_{\cdot 122}S_{\cdot 1}] = 3(r)_2\mu_{r-1}\mu_2 - (r)_3\mu_{r-3}\mu_2^2.
\end{aligned}$$

For example, the asymptotic variance of $n^{1/2}(T(\widehat{F}) - T(F))$ is

$$\begin{aligned}
&\mu^{-2r}\left(r^2\mu_{r-1}^2\mu_2 - 2r\mu_{r-1}\mu_{r+1} + \mu_{2r} - \mu_r^2\right)\\
&-2r\mu^{-2r-1}\left(\mu_{r+1} - r\mu_{r-1}\mu_2\right) + r^2\mu^{-2r-2}\mu_r^2\left(\mu_4 - \mu_2^2\right).
\end{aligned}$$

For $r = 2$, this reduces to $T(F)^2(\mu_4\mu_2^{-2} - 1 - 4\mu_3\mu_2^{-1}\mu^{-1} + 4\mu_2\mu^{-2})$.

Example 4.5 *This example is about the coefficient of variation. Suppose $T(F) = \mu_2^{1/2}/\mu$. Then the $[\cdot]_T$ needed for third order inference and bias reduction are given by Example 3.4 with $g_j = (1/2)_j\mu_2^{1/2-j}$, $S = \mu_2$, $U = \mu^{-1}$. By Example 4.1, $U_{\cdot 1\cdots p} = (-1)_p\mu^{-1-p}h_1\cdots h_p$*

so the terms needed in Example 3.4 are

$$
\begin{aligned}
&\left[1^2\right]_U = \mu^{-4}\mu_2,\ [U_{\cdot 1}S_{\cdot 1}] = -\mu^{-2}\mu_3,\ \left[1^2\right]_S = \mu_4 - \mu_2^2,\\
&[11]_U = 2\mu^{-3}\mu_2,\ [11]_S = -2\mu_2,\ \left[1^3\right]_U = -\mu^{-6}\mu_3,\\
&\left[U_{\cdot 1}^2 S_{\cdot 1}\right] = \mu^{-4}\left(\mu_4 - \mu_2^2\right),\ \left[U_{\cdot 1}S_{\cdot 1}^2\right] = -\mu^{-2}\left(\mu_5 - 2\mu_3\mu_2\right),\\
&\left[1^3\right]_S = \mu_6 - 3\mu_4\mu_2 + 2\mu_2^3,\ [1,2,12]_U = 2\mu^{-7}\mu_2^2,\\
&[U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}] = -2\mu^{-4}\mu_2^2,\ [U_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = -2\mu^{-5}\mu_3\mu_2,\\
&[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] = 2\mu^{-2}\mu_3\mu_2,\\
&[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] = 2\mu^{-3}\mu_3^2,\ [1,2,12]_S = -2\mu_3^2,\\
&[111]_U = -6\mu^{-4}\mu_3,\ [111]_S = 0,\ [U_{\cdot 1}S_{\cdot 11}] = 2\mu^{-2}\mu_3,\\
&[U_{\cdot 11}S_{\cdot 1}] = 2\mu^{-3}\left(\mu_4 - \mu_2^2\right),\ [S_{\cdot 1}S_{\cdot 11}] = -2\left(\mu_4 - \mu_2^2\right),\\
&[1122]_U = 24\mu^{-5}\mu_2^2,\ [1122]_S = 0,\ [U_{\cdot 1}S_{\cdot 122}] = 0,\\
&[U_{\cdot 122}S_{\cdot 1}] = -6\mu^{-4}\mu_2^2,\ [U_{\cdot 12}S_{\cdot 12}] = -4\mu^{-3}\mu_2^2,\\
&[1,122]_S = 0,\ \left[11^2\right]_S = 4\mu_4,\ \left[12^2\right]_S = 4\mu_2^2,\\
&[U_{\cdot 1}S_{\cdot 1}S_{\cdot 12}] = 2\mu^{-2}\mu_3\mu_2.
\end{aligned}
$$

For example,

$$a_{21} = \left[1^2\right] = T(F)^2\left(\mu_2\mu^{-2} - \mu_3\mu^{-1}\mu_2^{-1} + \mu_4\mu_2^{-2}4^{-1} - 4^{-1}\right)$$

as given by Section 10.6 of Stuart and Ord (1987). Also

$$
\begin{aligned}
&a_{11} = [11]/2 = T(F)\left(-3/8 - \mu_2^{-2}\mu_4/8 - \mu^{-1}\mu_2^{-1}\mu_3/2 + \mu^{-2}\mu_2\right),\\
&\left[1^3\right] = T(F)^3 \sum_{i=0}^{3} \mu^{-i}A_i,\\
&[1,2,12] = T(F)^3 \sum_{i=0}^{4} \mu^{-i}B_i,\\
&a_{32} = T(F)^3 \sum_{i=0}^{4} \mu^{-i}C_i,
\end{aligned}
$$

where

$$
\begin{aligned}
&A_0 = 1/4 - 3\mu_2^{-2}\mu_4/8 + \mu_2^{-3}\mu_6/8,\\
&A_1 = 3\left(2\mu_2^{-1}\mu_3 - \mu_2^{-2}\mu_5\right)/4,\ A_2 = 3\left(-\mu_2 + \mu_2^{-1}\mu_4\right)/2,\ A_3 = -\mu_3,\\
&B_0 = -\left(1 - \mu_2^{-2}\mu_4\right)^2/16 - \mu_2^{-2}\mu_3^2/4,\\
&B_1 = \mu_2^{-1}\mu_3/2,\ B_2 = -3\mu_2/2 + \mu_2^{-1}\mu_4/2 + 3\mu_2^{-2}\mu_3^2/4,\ B_3 = -3\mu_3,\\
&B_4 = 2\mu_2^2,\\
&C_0 = 1/16 - 3\mu_2^{-2}\mu_3^2/4 + \mu_2^{-3}\mu_6/8 - 3\mu_2^{-4}\mu_4^2/16,\\
&C_1 = 3\left(\mu_2^{-1}\mu_3 - \mu_2^{-2}\mu_5/4\right),\ C_2 = 3\left(-\mu_2 + \mu_2^{-1}\mu_4 + 3\mu_2^{-2}\mu_3^2/4\right),\\
&C_3 = -10\mu_3,\ C_4 = 6\mu_2^2.
\end{aligned}
$$

Appendix A

Continuing Example 3.4, the terms needed for (3.8)-(3.9) are:

$$
\begin{aligned}
[1,11]_T &= g_0^2\,[1,11]_U + g_0 g_1 \Bigg(U\,[S_{\cdot 1} U_{\cdot 11}] + U\,[U_{\cdot 1} S_{\cdot 11}] \\
&\quad +2\left[U_{\cdot 1}^2 S_{\cdot 1}\right] \Bigg) + g_0 g_2 U \left[U_{\cdot 1} S_{\cdot 1}^2\right] + g_1 g_2 U^2 \left[1^3\right]_S, \\
\left[12^2\right]_T &= g_0^2 \left[12^2\right]_U + 2 g_0 g_1 \Bigg(U\,[U_{\cdot 12} S_{\cdot 12}] + 2\,[U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] \Bigg) \\
&\quad + g_1^2 \Bigg(U^2 \left[12^2\right]_S + 4U\,[U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}] + 2\left[1^2\right]_U \left[1^2\right]_S + 2\,[U_{\cdot 1} S_{\cdot 1}]^2 \Bigg) \\
&\quad + 2 g_0 g_2 U\,[S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] + 2 g_1 g_2 U \Bigg(U\,[1,2,12]_S + 2\,[U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S \Bigg) \\
&\quad + g_2^2 U^2 \left[1^2\right]_S^2 \Bigg), \\
[1,122]_T &= g_0^2\,[1,122]_U + g_0 g_1 \Bigg(U\,[U_{\cdot 1} S_{\cdot 122}] + \left[1^2\right]_U [11]_S + 2\,[U_{\cdot 1} U_{\cdot 2} S_{\cdot 12}] \\
&\quad + 2\,[U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] + [U_{\cdot 1} S_{\cdot 1}]\,[11]_U + U\,[S_{\cdot 1} U_{\cdot 122}] \Bigg) \\
&\quad + g_0 g_2 \Bigg(U\,[U_{\cdot 1} S_{\cdot 1}]\,[11]_S + 2U\,[U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}] + \left[1^2\right]_U \left[1^2\right]_S + 2\,[U_{\cdot 1} S_{\cdot 1}]^2 \Bigg) \\
&\quad + g_0 g_3 U\,[U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S + g_1^2 U \Bigg(U\,[1,122]_S + [S_{\cdot 1} U_{\cdot 1}]\,[11]_S \\
&\quad + 2\,[U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}] + \left[1^2\right]_S [11]_U + 2\,[S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] \Bigg) + g_1 g_2 U \Bigg(U \left[1^2\right]_S [11]_S \\
&\quad + 2U\,[1,2,12]_S + [S_{\cdot 1} U_{\cdot 1}] \left[1^2\right]_S + 2\left[1^2\right]_S [U_{\cdot 1} S_{\cdot 1}] \Bigg) + 2 g_1 g_3 U^2 \left[1^2\right]_S^2 \Bigg), \\
\left[1^4\right]_T &= g_0^4 \left[1^4\right]_U + 4 g_0^3 g_1 U \left[U_{\cdot 1}^3 S_{\cdot 1}\right] + 6 g_0^2 g_1^2 U^2 \left[U_{\cdot 1}^2 S_{\cdot 1}^2\right] \\
&\quad + 4 g_0 g_1^3 U^3 \left[U_{\cdot 1} S_{\cdot 1}^3\right] + g_1^4 U^4 \left[1^4\right]_S,
\end{aligned}
$$

$$
\begin{aligned}
\left[1,2^2,12\right]_T \;=\;& g_0^4\left[1,2^2,12\right]_U + g_0^3 g_1\Bigg(U\left[U_{\cdot 1}U_{\cdot 2}^2 S_{\cdot 12}\right] + \left[1^2\right]_U\left[U_{\cdot 1}^2 S_{\cdot 1}\right] \\
& + \left[1^3\right]_U\left[U_{\cdot 1}S_{\cdot 1}\right] + 2U\left[U_{\cdot 1}U_{\cdot 2}S_{\cdot 2}U_{\cdot 12}\right] + U\left[U_{\cdot 1}^2 S_{\cdot 2}U_{\cdot 12}\right]\Bigg) \\
& + g_0^3 g_2 U\left[U_{\cdot 1}S_{\cdot 1}\right]\left[U_{\cdot 1}^2 S_{\cdot 1}\right] \\
& + g_0^2 g_1^2 U\Bigg(2U\left[U_{\cdot 1}U_{\cdot 2}S_{\cdot 2}S_{\cdot 12}\right] + 2\left[1^2\right]_U\left[U_{\cdot 1}S_{\cdot 1}^2\right] + 2\left[U_{\cdot 1}S_{\cdot 1}\right]\left[U_{\cdot 1}^2 S_{\cdot 1}\right] \\
& + U\left[U_{\cdot 1}S_{\cdot 2}^2 U_{\cdot 12}\right] + U\left[U_{\cdot 1}^2 S_{\cdot 2}S_{\cdot 12}\right] + \left[U_{\cdot 1}S_{\cdot 1}\right]\left[U_{\cdot 1}^2 S_{\cdot 1}\right] + \left[1^3\right]_U\left[1^2\right]_S\Bigg) \\
& + g_0^2 g_1 g_2 U^2\Bigg(2\left[U_{\cdot 1}S_{\cdot 1}\right]\left[U_{\cdot 1}S_{\cdot 1}^2\right] + \left[U_{\cdot 1}^2 S_{\cdot 1}\right]\left[1^2\right]_S\Bigg) \\
& + g_0 g_1^3 U^2\Bigg(U\left[U_{\cdot 1}S_{\cdot 2}^2 U_{\cdot 12}\right] + \left[1^2\right]_U\left[1^3\right]_S + \left[U_{\cdot 1}S_{\cdot 1}\right]\left[U_{\cdot 1}S_{\cdot 1}^2\right] + 2U\left[S_{\cdot 1}S_{\cdot 2}^2 U_{\cdot 12}\right]\Bigg) \\
& + g_0 g_1^2 g_2 U^3\left[U_{\cdot 1}S_{\cdot 1}\right]\left[1^3\right]_S \\
& + 2g_1^4 U^3\Bigg(U\left[1,2^2,12\right]_S + \left[U_{\cdot 1}S_{\cdot 1}\right]\left[1^3\right]_S + \left[U_{\cdot 1}S_{\cdot 1}^2\right]\left[1^2\right]_S\Bigg) \\
& + 2g_1^3 g_2 U^4\left[1^2\right]_S\left[1^3\right]_S\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
[1,2,13,23]_T \quad = \quad & g_0^4 [1,2,13,23]_U \\
& + g_0^3 g_1 \Bigg(2U [U_{\cdot 1} S_{\cdot 2} U_{\cdot 13} U_{\cdot 23}] + 2U [U_{\cdot 1} U_{\cdot 2} U_{\cdot 13} S_{\cdot 23}] + 2 [U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] \left[U_{\cdot 1}^2\right] \\
& + 2 [U_{\cdot 1} S_{\cdot 1}] [1,2,12]_U \Bigg) \\
& + g_0^3 g_1^2 \Bigg(U^2 [S_{\cdot 1} S_{\cdot 2} U_{\cdot 13} U_{\cdot 23}] + 2U^2 [U_{\cdot 1} S_{\cdot 2} U_{\cdot 23} S_{\cdot 13}] + 2U^2 [U_{\cdot 1} S_{\cdot 2} U_{\cdot 13} S_{\cdot 23}] \\
& + 2U \left[1^2\right]_U [S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] + 4U [U_{\cdot 1} S_{\cdot 1}] [U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] + 2U [1,2,12]_U \left[1^2\right]_S \\
& + U^2 [U_{\cdot 1} U_{\cdot 2} S_{\cdot 13} S_{\cdot 23}] + 2U [U_{\cdot 1} S_{\cdot 12} S_{\cdot 2}] \left[1^2\right]_U + 2U [U_{\cdot 1} U_{\cdot 2} S_{\cdot 12}] [U_{\cdot 1} S_{\cdot 1}] \\
& + \left[1^2\right]_U^2 \left[1^2\right]_S + 3 [U_{\cdot 1} S_{\cdot 1}]^2 \left[1^2\right]_S \Bigg) \\
& + 2 g_0 g_1 U \Bigg(U^2 [S_{\cdot 1} S_{\cdot 2} U_{\cdot 13} S_{\cdot 23}] + U [U_{\cdot 1} S_{\cdot 1}] [S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] U [U_{\cdot 1} U_{\cdot 12} S_{\cdot 2}] \left[1^2\right]_S \\
& + U^2 [U_{\cdot 1} S_{\cdot 2} S_{\cdot 13} S_{\cdot 23}] + U \left[1^2\right]_U [1, S_{\cdot 2}, 12]_S + 2U [U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}] [U_{\cdot 1} S_{\cdot 1}] \\
& + U [U_{\cdot 1} U_{\cdot 2} S_{\cdot 12}] \left[1^2\right]_S + 3 \left[1^2\right]_U [U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S + [U_{\cdot 1} S_{\cdot 1}]^3 \Bigg) \\
& + 2 g_0^3 g_2 U [U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] [U_{\cdot 1} S_{\cdot 1}] \\
& + 2 g_0^2 g_1 g_2 \Bigg(U^2 [U_{\cdot 1} S_{\cdot 1}] [S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] + U^2 [U_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] \left[1^2\right]_S + U^2 [U_{\cdot 1} S_{\cdot 2} S_{\cdot 12}] [U_{\cdot 1} S_{\cdot 1}] \\
& + U \left[1^2\right]_U [U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S + U [U_{\cdot 1} S_{\cdot 1}]^3 \Bigg) \\
& + 2 g_0 g_1^2 g_2 U^2 \Bigg(U [S_{\cdot 1} S_{\cdot 2} U_{\cdot 12}] \left[1^2\right]_S + U [U_{\cdot 1} S_{\cdot 1}] [1,2,12]_S + U [U_{\cdot 1} S_{\cdot 1} S_{\cdot 12}] \left[1^2\right]_S \\
& + 3 [U_{\cdot 1} S_{\cdot 1}]^2 \left[1^2\right]_S + \left[1^2\right]_U \left[1^2\right]_S^2 \Bigg) \\
& + 2 g_1^3 g_2 U^3 \Bigg(U \left[1^2\right]_S [1,2,12]_S + 2 [U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S^2 \Bigg) \\
& + g_0^2 g_2^2 U^2 [U_{\cdot 1} S_{\cdot 1}]^2 \left[1^2\right]_S \\
& + 2 g_0 g_1 g_2^2 U^3 [U_{\cdot 1} S_{\cdot 1}] \left[1^2\right]_S^2 \\
& + g_1^2 g_2^2 U^4 \left[1^2\right]_S^3 \Bigg),
\end{aligned}
$$

$$
\begin{aligned}
[1,2,3,123]_T &= g_0^4\,[1,2,3,123]_U \\
&+g_0^3 g_1\Bigg(U\,[U_{\cdot 1}U_{\cdot 2}U_{\cdot 3}S_{\cdot 123}] + 3\left[1^2\right]_U [U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}] + 3\,[1,2,12]_U\,[U_{\cdot 1}S_{\cdot 1}] \\
&+3U\,[U_{\cdot 1}U_{\cdot 2}S_{\cdot 3}U_{\cdot 123}]\Bigg) \\
&+g_0^3 g_2\Bigg(3U\,[U_{\cdot 1}S_{\cdot 1}]\,[U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}] + 3\left[1^2\right]_U [U_{\cdot 1}S_{\cdot 1}]^2\Bigg) \\
&+g_0^3 g_3 U\,[U_{\cdot 1}S_{\cdot 1}]^3 \\
&+3g_0^2 g_1^2 U\Bigg(U\,[U_{\cdot 1}U_{\cdot 2}S_{\cdot 3}S_{\cdot 123}] + 2\left[1^2\right]_U [U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] + [U_{\cdot 1}S_{\cdot 1}]\,[U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}] \\
&+2\,[U_{\cdot 1}S_{\cdot 1}]\,[U_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] + [1,2,12]_U\left[1^2\right]_S + U\,[U_{\cdot 1}S_{\cdot 1}S_{\cdot 3}U_{\cdot 123}]\Bigg) \\
&+3g_0^2 g_1 g_2 U\Bigg(2U\,[U_{\cdot 1}S_{\cdot 1}]\,[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] + U\,[U_{\cdot 1}U_{\cdot 2}S_{\cdot 12}]\left[1^2\right]_S \\
&+2\left[1^2\right]_U [U_{\cdot 1}S_{\cdot 1}]\left[1^2\right]_S + [U_{\cdot 1}S_{\cdot 1}]^3\Bigg) \\
&+3g_0^2 g_1 g_3 U^2\,[U_{\cdot 1}S_{\cdot 1}]^2\left[1^2\right]_S \\
&+g_0 g_1^3 U^2\Bigg(3U\,[U_{\cdot 1}S_{\cdot 2}S_{\cdot 3}S_{\cdot 123}] + 3\left[1^2\right]_U [S_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] + 6\,[U_{\cdot 1}S_{\cdot 1}]\,[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}] \\
&+3\,[U_{\cdot 1}S_{\cdot 1}]\,[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}] + 6\,[U_{\cdot 1}S_{\cdot 2}U_{\cdot 12}]\left[S_{\cdot 1}^2\right] + U\,[S_{\cdot 1}S_{\cdot 2}S_{\cdot 3}U_{\cdot 123}]\Bigg) \\
&+3g_0 g_1^2 g_2 U^2\Bigg(U\,[U_{\cdot 1}S_{\cdot 1}]\,[1,2,12]_S + 2U\,[U_{\cdot 1}S_{\cdot 2}S_{\cdot 12}]\left[1^2\right]_S \\
&+\left[1^2\right]_U\left[1^2\right]_S^2 + 2\,[U_{\cdot 1}S_{\cdot 1}]^2\left[1^2\right]_S\Bigg) \\
&+3g_0 g_1^2 g_3 U^3\,[U_{\cdot 1}S_{\cdot 1}]\left[1^2\right]_S^2 \\
&+g_1^4 U^3\Bigg(U\,[1,2,3,123]_S + 3\,[U_{\cdot 1}S_{\cdot 1}]\,[1,2,12]_S + 3\,[S_{\cdot 1}S_{\cdot 2}U_{\cdot 12}]\left[1^2\right]_S\Bigg) \\
&+3g_1^3 g_2 U^3\Bigg(U\,[1]_S^2\,[1,2,12]_S + [U_{\cdot 1}S_{\cdot 1}]\left[1^2\right]_S^2\Bigg) \\
&+g_1^3 g_3 U^4\left[1^2\right]_S^3 .
\end{aligned}
$$

Appendix B

Here, we give bracket functions for moments to order seven, that is expressions for

$$\left(1_r^a 11_s^b 2_t^2 12_u^d \cdots\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mu_{r\cdot 1}^a \mu_{s\cdot 11}^b \mu_{t\cdot 2}^c \mu_{u\cdot 12}^d \cdots dF_1(x_1)\, dF_2(x_2) \cdots$$

up to $N = ar + bs + ct + \cdots = 7$, where now μ_1 means μ. This enables one to obtain the Edgeworth-Cornish-Fisher expansions and bias reduction for any smooth function of $(\mu, \mu_2, \mu_3, \ldots)$. We exclude separable terms like $(1_a^2 2_b^2) = (1_a^2)(1_b^2)$. For each N, we order the terms by its partition functions.

$$\begin{array}{rcl}
N & = & 2 \text{ (2 terms)}: \\
1^2: & & \left(1_1^2\right) = \mu_2, \\
2: & & (11_2) = -2\mu_2. \\
N & = & 3 \text{ (5 terms)}: \\
1^3: & & \left(1_1^3\right) = \mu_3, \\
12: & & (1_1 1_2) = \mu_3,\ (1_1 11_2) = -2\mu_3, \\
3: & & (11_3) = -6\mu_3,\ (111_3) = -12\mu_3. \\
N & = & 4 \text{ (13 terms)}: \\
1^4: & & \left(1_1^4\right) = \mu_4, \\
1^2 2: & & \left(1_1^2 1_2\right) = \mu_4 - \mu_2^2,\ \left(1_1^2 11_2\right) = -2\mu_4, \\
13: & & (1_1 1_3) = \mu_4 - \mu_2^2,\ (1_1 11_3) = -6\left(\mu_4 - \mu_2^2\right),\ (1_1 111_3) = 12\mu_4, \\
& & (1_1 122_3) = 12\mu_2^2, \\
2^2: & & \left(1_2^2\right) = \mu_4 - \mu_2^2,\ (1_2 11_2) = -2\left(\mu_4 - \mu_2^2\right),\ \left(11_2^2\right) = 4\mu_4,\ \left(12_2^2\right) = 4\mu_2^2, \\
4: & & (1111_4) = -72\mu_4,\ (1122_4) = -72\mu_2^2. \\
N & = & 5 \text{ (43 terms)}: \\
1^5: & & \left(1_1^5\right) = \mu_5, \\
1^3 2: & & \left(1_1^3 1_2\right) = \mu_5 - \mu_3\mu_2,\ \left(1_1^3 11_2\right) = \mu_5 - 2\mu_3\mu_2,\ \left(1_1^2 2_1 12_2\right) = -2\mu_3\mu_2, \\
1^2 3: & & \left(1_1^2 1_3\right) = \mu_3\mu_2,\ \left(1_1^2 11_3\right) = -6\left(\mu_5 - \mu_3\mu_2\right),\ \left(1_1^2 111_3\right) = 12\mu_5, \\
& & \left(1_1^2 122_3\right) = 12\mu_3\mu_2,\ (1_1 2_1 12_3) = -2\mu_2^3,\ (1_1 2_1 112_3) = 12\mu_3\mu_2, \\
12^2: & & \left(1_1 1_2^2\right) = \mu_5 - 2\mu_3\mu_2,\ (1_1 1_2 11_2) = -2\left(\mu_5 - \mu_3\mu_2\right),\ (1_1 2_2 12_2) = 4\mu_3\mu_2, \\
& & \left(1_1 11_2^2\right) = 4\mu_5,\ \left(1_1 12_2^2\right) = 4\mu_3\mu_2, \\
14: & & (1_1 1_4) = \mu_5 - 4\mu_3\mu_2,\ (1_1 11_4) = -8\mu_5 + 20\mu_3\mu_2, \\
& & (1_1 111_4) = 36\left(\mu_5 - \mu_3\mu_2\right),\ (1_1 1111_4) = -72\mu_5, \\
& & (1_1 122_4) = 6\mu_3\mu_2,\ (1_1 1122_4) = -72\mu_3\mu_2, \\
& & (1_1 1222_4) = -72\mu_3\mu_2, \\
23: & & (1_2 1_3) = \mu_5 - 4\mu_3\mu_2,\ (1_2 11_3) = -6\left(\mu_5 - \mu_3\mu_2\right),\ (1_2 111_3) = 12\left(\mu_5 - \mu_3\mu_2\right), \\
& & (11_2 1_3) = -2\left(\mu_5 - 4\mu_3\mu_2\right),\ (11_2 11_3) = 12\left(\mu_5 - \mu_3\mu_2\right),\ (11_2 111_3) = -24\mu_5, \\
& & (11_2 122_3) = -24\mu_3\mu_2,\ (12_2 112_3) = -24\mu_3\mu_2, \\
5: & & (11_5) = -10\left(\mu_5 - 2\mu_3\mu_2\right),\ (111_5) = 60\left(\mu_5 - 2\mu_3\mu_2\right),\ (1111_5) = -240\left(\mu_5 - \mu_3\mu_2\right), \\
& & (11111_5) = 480\mu_5,\ (1122_5) = 0,\ (11122_5) = 480\mu_3\mu_2.
\end{array}$$

$$
\begin{aligned}
N &= 6 \;\; (85 \text{ terms}): \\
1^6: &\quad \left(1_1^6\right)=\mu_6, \\
1^4 2: &\quad \left(1_1^4 1_2\right)=\mu_6-\mu_4\mu_2,\ \left(1_1^4 11_2\right)=-2\mu_6,\ \left(1_1^3 2_1 12_2\right)=-2\mu_4\mu_2, \\
&\quad \left(1_1^2 2_1^2 12_2\right)=-2\mu_3^2, \\
1^3 3: &\quad \left(1_1^3 1_3\right)=\mu_6-3\mu_4\mu_2-\mu_3^2,\ \left(1_1^3 11_3\right)=-6\left(\mu_6-\mu_4\mu_2\right), \\
&\quad \left(1_1^3 111_3\right)=12\mu_6,\ \left(1_1^3 122_3\right)=12\mu_4\mu_2,\ \left(1_1^2 2_1 12_3\right)=-3\left(\mu_4\mu_2+\mu_3^2-\mu_2^3\right), \\
&\quad \left(1_1^2 2_1 112_3\right)=12\mu_4\mu_2,\ \left(1_1^2 2_1 122_3\right)=12\mu_3^2,\ \left(1_1 2_1 3_1 123_3\right)=12\mu_2^3, \\
1^2 4: &\quad \left(1_1^2 1_4\right)=\mu_6-4\mu_4\mu_2-4\mu_3^2,\ \left(1_1^2 11_4\right)=-4\left(2\mu_6-3\mu_4\mu_2-2\mu_3^2\right), \\
&\quad \left(1_1^2 111_4\right)=-12\left(\mu_6-\mu_4\mu_2\right),\ \left(1_1^2 1111_4\right)=-72\mu_6, \\
&\quad \left(1_1^2 122_4\right)=-12\left(\mu_4\mu_2+2\mu_3^2-\mu_2^3\right),\ \left(1_1^2 1122_4\right)=-72\mu_4\mu_2, \\
&\quad \left(1_1^2 1222_4\right)=-72\mu_3^2,\ \left(1_1 2_1 12_4\right)=-4\left(2\mu_4\mu_2-3\mu_2^3\right), \\
&\quad \left(1_1 2_1 112_4\right)=-12\left(2\mu_4\mu_2+\mu_3^2-2\mu_2^3\right),\ \left(1_1 2_1 1112_4\right)=480\mu_4\mu_2, \\
&\quad \left(1_1 2_1 1122_4\right)=480\mu_3^2,\ \left(1_1 2_1 1233_4\right)=480\mu_2^3, \\
1^2 2^2: &\quad \left(1_1^2 11_2^2\right)=4\mu_6,\ \left(1_1^2 12_2^2\right)=4\mu_4\mu_2,\ \left(1_1 2_1 11_2 22_2\right)=4\mu_3^2, \\
&\quad \left(1_1 2_1 12_2^2\right)=-2\mu_3^2, \\
123: &\quad \left(1_1 11_2 111_3\right)=-24\mu_6,\ \left(1_1 11_2 122_3\right)=-24\mu_4\mu_2, \\
&\quad \left(1_1 12_2 112_3\right)=-24\mu_4\mu_2,\ \left(1_1 12_2 122_3\right)=-24\mu_3^2, \\
&\quad \left(1_1 12_2 222_3\right)=-24\mu_4\mu_2, \\
15: &\quad \left(1_1 1_5\right)=\mu_6-5\mu_4\mu_2,\ \left(1_1 11_5\right)=-10\left(\mu_6-\mu_4\mu_2-2\mu_3^2\right), \\
&\quad \left(1_1 111_5\right)=60\left(\mu_6-\mu_4\mu_2-\mu_3^2\right),\ \left(1_1 122_5\right)=60\mu_4\mu_2, \\
&\quad \left(1_1 1111_5\right)=240\left(\mu_6-\mu_4\mu_2\right),\ \left(1_1 1122_5\right)=120\left(\mu_4\mu_2+\mu_3^2-\mu_2^3\right), \\
&\quad \left(1_1 1222_5\right)=60\left(3\mu_4\mu_2+\mu_3^2-3\mu_2^3\right),\ \left(1_1 11111_5\right)=480\mu_6, \\
&\quad \left(1_1 11122_5\right)=480\mu_4\mu_2,\ \left(1_1 11222_5\right)=480\mu_3^2, \\
&\quad \left(1_1 12222_5\right)=480\mu_4\mu_2,\ \left(1_1 12233_5\right)=480\mu_2^3, \\
2^3: &\quad \left(1_2^3\right)=\mu_6-3\mu_4\mu_2+2\mu_2^3,\ \left(1_2 1_2 11_2\right)=-2\left(\mu_6-2\mu_4\mu_2+\mu_2^3\right), \\
&\quad \left(1_2 2_2 1_2\right)=2\mu_2^3,\ \left(1_2 11_2^2\right)=4\left(\mu_6-\mu_4\mu_2\right), \\
&\quad \left(1_2 12_2^2\right)=4\left(\mu_4\mu_2-\mu_2^3\right),\ \left(11_2^3\right)=-8\mu_6, \\
&\quad \left(11_2 12_2^2\right)=-8\mu_4\mu_2,\ \left(12_2 23_2 31_2\right)=-8\mu_2^3, \\
24: &\quad \left(1_2 1_4\right)=\mu_6-\mu_4\mu_2-4\mu_2^3,\ \left(1_2 11_4\right)=-4\left(2\mu_6-3\mu_4\mu_2-2\mu_3^2+3\mu_2^3\right), \\
&\quad \left(1_2 111_4\right)=-36\left(\mu_6-2\mu_4\mu_2+2\mu_2^3\right),\ \left(1_2 1111_4\right)=-72\left(\mu_6-\mu_4\mu_2\right), \\
&\quad \left(1_2 1122_4\right)=-72\left(\mu_4\mu_2-\mu_2^3\right),\ \left(11_2 1_4\right)=-2\left(\mu_6-\mu_4\mu_2-4\mu_2^3\right), \\
&\quad \left(11_2 11_4\right)=8\left(2\mu_6-3\mu_4\mu_2-2\mu_3^2\right),\ \left(11_2 111_4\right)=72\left(\mu_6-\mu_4\mu_2\right), \\
&\quad \left(11_2 1111_4\right)=144\mu_6,\ \left(11_2 1122_4\right)=144\mu_4\mu_2, \\
&\quad \left(12_2 12_4\right)=8\left(2\mu_4\mu_2-3\mu_2^3\right),\ \left(12_2 112_4\right)=-4\left(2\mu_4\mu_2+\mu_3^2-2\mu_2^3\right), \\
&\quad \left(12_2 1112_4\right)=\ \left(12_2 1122_4\right)=144\mu_3^2, \\
3^2: &\quad \left(1_3^2\right)=\mu_6-6\mu_4\mu_2-\mu_3^2+9\mu_2^3,\ \left(1_3 11_3\right)=-6\left(\mu_6-4\mu_4\mu_2-\mu_3^2+3\mu_2^3\right), \\
&\quad \left(1_3 111_3\right)=12\left(\mu_6-3\mu_4\mu_2-\mu_3^2\right),\ \left(1_3 122_3\right)=12\left(\mu_4\mu_2-3\mu_2^3\right), \\
&\quad \left(11_3 122_3\right)=-72\left(\mu_4\mu_2-\mu_2^3\right),\ \left(12_3^2\right)=18\left(\mu_4\mu_2+\mu_3^2-\mu_2^3\right), \\
&\quad \left(12_3 112_3\right)=-36\left(\mu_4\mu_2+\mu_3^2-\mu_2^3\right),
\end{aligned}
$$

$6:$

$$\begin{aligned}
&(11_6) = -6\left(2\mu_6 - 5\mu_4\mu_2\right),\ (111_6) = 30\left(3\mu_6 - 3\mu_4\mu_2 - 4\mu_3^2\right),\\
&(1111_6) = -120\left(\mu_6 - \mu_4\mu_2 + \mu_3^2\right),\ (11111_6) = -1800\left(\mu_6 - \mu_4\mu_2\right),\\
&(1122_6) = -120\left(4\mu_4\mu_2 - 3\mu_2^3\right),\ (11122_6) = -360\left(3\mu_4\mu_2 + 2\mu_3^2 - 3\mu_2^3\right),\\
&(111111_6) = -3600\mu_6,\ (111122_6) = -3600\mu_4\mu_2,\\
&(111222_6) = -3600\mu_3^2,\ (112233_6) = -3600\mu_2^3.
\end{aligned}$$

As noted for $N = 2, \ldots, 6$ there are 2,5,13,43 and 85 terms, or without μ, 1,2,6,20 and 39 terms. We now give the 90 terms for $N = 7$ without μ.

$$
\begin{aligned}
N &= 7: \\
2^2 3: \quad & \left(1_2^2 1_3\right) = \mu_5 - 4\mu_3\mu_2, \\
& \left(1_2^2 11_3\right) = -6\left(\mu_7 - 3\mu_5\mu_2 + 3\mu_3\mu_2^2\right), \\
& \left(1_2^2 111_3\right) = -24\mu_5, \\
& \left(1_2^2 122_3\right) = \left(1_2 2_2 122_3\right) = -24\mu_3\mu_2, \\
& \left(1_2 11_2 1_3\right) = 2\left(3\mu_5\mu_2 - \mu_4\mu_3 + \mu_3^2\right), \\
& \left(1_2 11_2 11_3\right) = 12\left(\mu_7 - 2\mu_5\mu_2 + \mu_3\mu_2^2\right), \\
& \left(1_2 11_2 111_3\right) = -24\left(\mu_7 - \mu_5\mu_2\right), \\
& \left(1_2 11_2 122_3\right) = -12\left(\mu_5 - \mu_3\mu_2\right)\mu_2, \\
& \left(1_2 12_2 2_3\right) = 4\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& \left(1_2 12_2 12_3\right) = 6\left(\mu_5\mu_2 - 2\mu_4\mu_3\right), \\
& \left(1_2 12_2 112_3\right) = -24\left(\mu_5 - \mu_3\mu_2\right)\mu_2, \\
& \left(1_2 12_2 122_3\right) = -24\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& \left(1_2 12_2 222_3\right) = -24\mu_4\mu_3, \\
& \left(1_2 12_2 233_3\right) = -24\mu_3\mu_2^2, \\
25: \quad & \left(1_2 1_5\right) = \mu_7 - \mu_5\mu_2 - 5\mu_4\mu_3, \\
& \left(1_2 11_5\right) = -10\left(\mu_7 - \mu_5\mu_2 + 3\mu_4\mu_3 - 2\mu_3\mu_2^2\right), \\
& \left(1_2 111_5\right) = 60\left(\mu_7 - 2\mu_5\mu_2 - \mu_4\mu_3 + 2\mu_3\mu_2^2\right), \\
& \left(1_2 122_5\right) = 20\left(\mu_5\mu_2 + 2\mu_4\mu_3 - 4\mu_3\mu_2^2\right), \\
& \left(1_2 1111_5\right) = \left(1_2 11111_5\right) = 480\left(\mu_7 - \mu_5\mu_2\right), \\
& \left(1_2 1122_5\right) = 120\left(\mu_5\mu_2 + \mu_4\mu_3 - 3\mu_3\mu_2^2\right), \\
& \left(1_2 11122_5\right) = \left(11_2 11122_5\right) = 480\left(\mu_5 - \mu_3\mu_2\right)\mu_2, \\
& \left(1_2 11222_5\right) = \left(11_2 11222_5\right) = -\left(12_2 1122_5\right) = 480\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& \left(1_2 12222_5\right) = \left(11_2 12222_5\right) = 480\mu_4\mu_3, \\
& \left(1_2 12233_5\right) = \left(11_2 12233_5\right) = 480\mu_3\mu_2^2, \\
& \left(11_2 1_5\right) = -2\left(\mu_7 - \mu_5\mu_2 - 5\mu_4\mu_3\right), \\
& \left(11_2 11_5\right) = 20\left(\mu_7 - 3\mu_4\mu_3\right), \\
& \left(11_2 111_5\right) = -120\left(\mu_7 - \mu_5\mu_2 - \mu_4\mu_3\right), \\
& \left(11_2 122_5\right) = -40\left(\mu_5\mu_2 + 2\mu_4\mu_3 - 4\mu_3\mu_2^2\right), \\
& \left(11_2 1111_5\right) = 240\left(\mu_7 - 2\mu_5\mu_2 + \mu_3\mu_2^2\right), \\
& \left(11_2 1122_5\right) = 120\left(\mu_5\mu_2 + \mu_4\mu_3 - 3\mu_3\mu_2^2\right), \\
& \left(11_2 11111_5\right) = 480\left(\mu_7 - \mu_5\mu_2\right), \\
& \left(12_2 112_5\right) = -40\left(2\mu_5\mu_2 + \mu_4\mu_3 - 5\mu_3\mu_2^2\right), \\
& \left(12_2 1112_5\right) = -120\left(3\mu_5\mu_2 + \mu_4\mu_3 - 3\mu_3\mu_2^2\right), \\
& \left(12_2 1222_5\right) = -120\left(3\mu_5\mu_2 + \mu_4\mu_3 - 3\mu_3\mu_2^2\right), \\
& \left(12_2 11112_5\right) = -960\mu_5\mu_2, \\
& \left(12_2 11122_5\right) = -960\mu_4\mu_3, \\
& \left(12_2 12233_5\right) = -960\mu_3\mu_2^2,
\end{aligned}
$$

$$
\begin{aligned}
34: \quad & (1_3 1_4) = \mu_7 - 3\mu_5\mu_2 - 5\mu_4\mu_3 + 12\mu_3\mu_2^2, \\
& (1_3 11_4) = -4\left(2\mu_7 - 9\mu_5\mu_2 - 4\mu_4\mu_3 - 18\mu_3\mu_2^2\right), \\
& (1_3 111_4) = -36\left(\mu_7 - 4\mu_5\mu_2 - \mu_4\mu_3 + 4\mu_3\mu_2^2\right), \\
& (1_3 1111_4) = 72\left(\mu_7 - 3\mu_5\mu_2 - \mu_4\mu_3\right), \\
& (1_3 1122_4) = 72\left(\mu_5 - 4\mu_3\mu_2\right)\mu_2, \\
& (11_3 1_4) = \ (11_3 11_4)/4 = -6\left(\mu_7 - \mu_5\mu_2 - 4\mu_4\mu_3 + 4\mu_3\mu_2^2\right), \\
& (11_3 111_4) = 216\left(\mu_7 - 2\mu_5\mu_2 + \mu_3\mu_2^2\right), \\
& (11_3 1111_4) = 432\left(\mu_7 - \mu_5\mu_2\right), \\
& (11_3 1122_4) = 432\left(\mu_5 - \mu_3\mu_2\right)\mu_2, \\
& (11_3 11222_4) = 432\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& (12_3 12_4) = 24\left(\mu_5\mu_2 + \mu_4\mu_3 - 5\mu_3\mu_2^2\right), \\
& (12_3 112_4) = 24\left(\mu_5\mu_2 + \mu_4\mu_3 - 4\mu_3\mu_2^2\right), \\
& (12_3 1112_4) = \ (12_3 1222_4) = 216\left(\mu_5\mu_2 + \mu_4\mu_3 - \mu_3\mu_2^2\right), \\
& (12_3 1122_4) = 432\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& (12_3 1233_4) = 432\mu_3\mu_2^2, \\
7: \quad & (11_7) = -14\left(\mu_7 - 3\mu_5\mu_2\right), \\
& (111_7) = 42\left(3\mu_7 - 3\mu_5\mu_2 - 5\mu_4\mu_3\right), \\
& (1111_7) = -840\left(\mu_7 - 2\mu_4\mu_3\right), \\
& (1122_7) = -140\left(4\mu_5\mu_2 - 3\mu_3\mu_2^2\right), \\
& (11111_7) = 840\left(5\mu_7 - 3\mu_5\mu_2 - 5\mu_4\mu_3\right), \\
& (11122_7) = 420\left(3\mu_5\mu_2 + 2\mu_4\mu_3 - 6\mu_3\mu_2^2\right), \\
& (111111_7) = 15120\left(\mu_7 - \mu_5\mu_2\right), \\
& (111122_7) = 7!\left(2\mu_5\mu_2 + \mu_4\mu_3 - 2\mu_3\mu_2^2\right), \\
& (111222_7) = 15120\left(\mu_4 - \mu_2^2\right)\mu_3, \\
& (112233_7) = 15120\mu_3\mu_2^2, \\
& (1111111_7) = 30240\mu_7, \\
& (1111122_7) = 30240\mu_5\mu_2, \\
& (1111222_7) = 30240\mu_4\mu_3, \\
& (1112233_7) = 30240\mu_3\mu_2^2.
\end{aligned}
$$

For F symmetric terms corresponding to odd N reduce to zero.

References

[1] Comtet, L. (1974). *Advanced Combinatorics*. Reidel, Dordrecht.

[2] D'Agostino, R. and Pearson, E. S. (1973). Tests for departure from normality, empirical results for the distribution of b_2 and $\sqrt{b_1}$. *Biometrika*, **60**, 613-622.

[3] Fisher, R. A. (1931). The moments of the distribution for normal samples of measures of departure from normality. *Proceedings of the Royal Society London*, A, **130**, 16-28.

[4] Loisel, S., Mazza, C. and Rulliere, D. (2009). Convergence and asymptotic variance of bootstrapped finite-time ruin probabilities with partly shifted risk processes. *Insurance: Mathematics and Economics*, **45**, 374-381.

[5] Nickl, R. (2007). Donsker-type theorems for nonparametric maximum likelihood estimators. *Probability Theory and Related Fields*, **138**, 411-449.

[6] Stuart, A. and Ord, J. K. (1987). *Kendall's Advanced Theory of Statistics*, volume 1, fifth edition. Griffin, London.

[7] Volodko, N. V. (2011). Limit theorems for canonical von Mises and U-statistics of m-dependent observations. *Theory of Probability and Its Applications*, **55**, 271-290.

[8] von Mises, R. (1947). Differentiable statistical functions. *Annals of Mathematical Statistics*, **18**, 309-348.

[9] Withers, C. S. (1982). Second order inference for asymptotically normal random variables. *Sankhyā*, B, **44**, 19-27.

[10] Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Annals of Statistics*, **11**, 577-587.

[11] Withers, C. S. (1984). A chain rule for differentiation with applications to multivariate Hermite polynomials. *Bulletin of the Australian Mathematical Society*, **30**, 247-250.

[12] Withers, C. S. (1987). Bias reduction by Taylor series. *Communications in Statistics—Theory and Methods*, **16**, 2369-2384.

[13] Withers, C. S. (1989). The distribution and cumulants of a Studentised statistic. *Communications in Statistics—Theory and Methods*, **16**, 295-318.

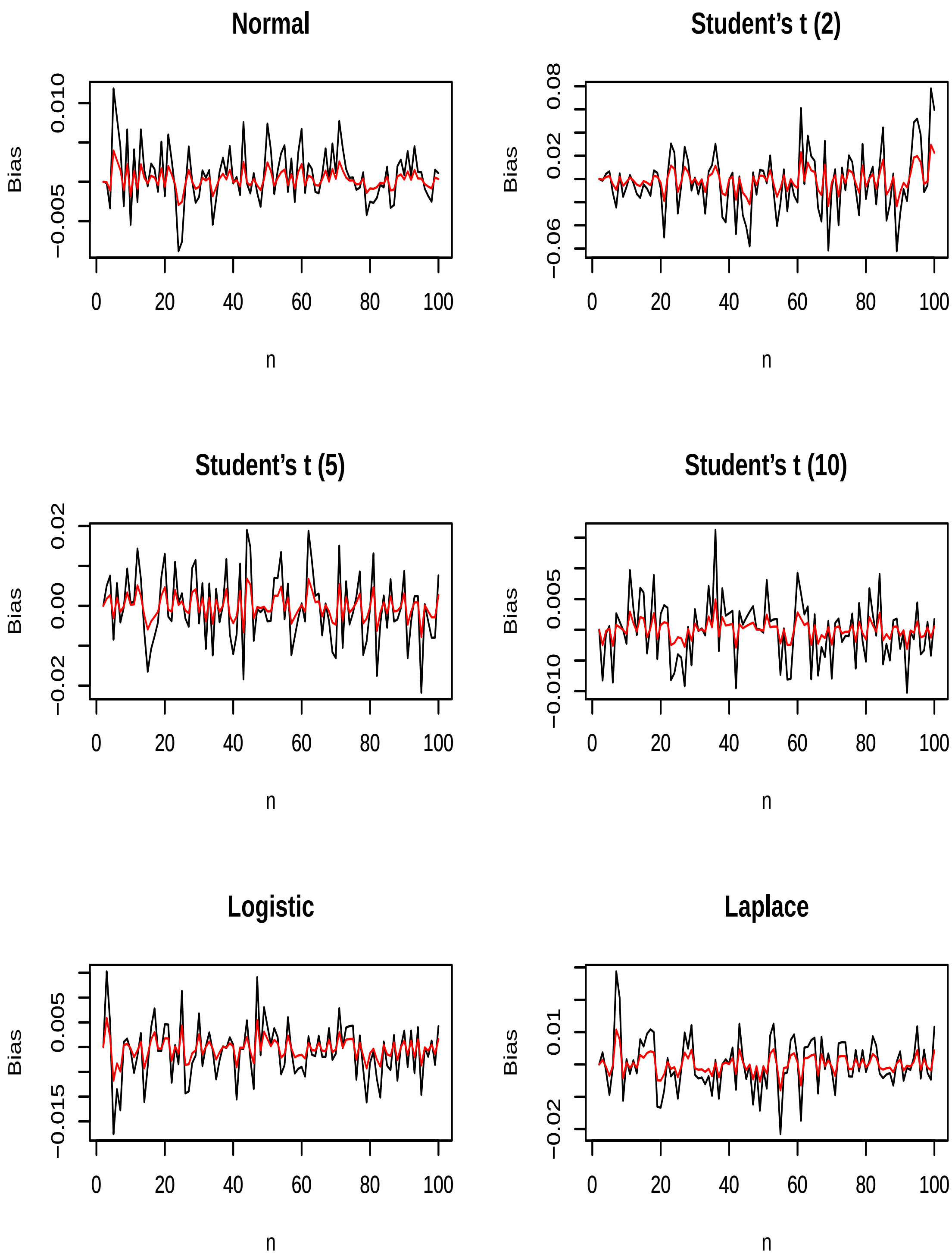

Figure 4.1 Biases of the usual (black) and bias reduced (red) estimators of skewness versus $n = 2, 3, \ldots, 100$.